University of Tennessee, Knoxville
Trace: Tennessee Research and Creative Exchange

Masters Theses | Graduate School

12-2003

Monte Carlo Simulations and Analysis of Single-Molecule Detection and Imaging

Peter Williams
University of Tennessee - Knoxville

Recommended Citation

Williams, Peter, "Monte Carlo Simulations and Analysis of Single-Molecule Detection and Imaging. " Master's Thesis, University of Tennessee, 2003.
https://trace.tennessee.edu/utk_gradthes/2337

This Thesis is brought to you for free and open access by the Graduate School at Trace: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Masters Theses by an authorized administrator of Trace: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.

To the Graduate Council:

I am submitting herewith a thesis written by Peter Williams entitled "Monte Carlo Simulations and Analysis of Single-Molecule Detection and Imaging." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Physics.

Dr. Lloyd M. Davis, Major Professor

We have read this thesis and recommend its acceptance:

Dr. Christian G. Parigger, Dr. Horace W. Crater

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a thesis written by Peter Williams entitled "Monte Carlo Simulations and Analysis of Single-Molecule Detection and Imaging." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Physics.

Dr. Lloyd M. Davis
Major Professor

We have read this thesis
and recommend its acceptance:

Dr. Christian G. Parigger

Dr. Horace W. Crater

Accepted for the Council:

Dr. Anne Mayhew
Vice Provost and Dean of Graduate Studies

(Original signatures are on file with official student records.)

Monte Carlo Simulations and Analysis of Single-Molecule Detection and Imaging

A Thesis Presented For the

Master of Science Degree

The University of Tennessee, Knoxville

Peter Edward Williams

December 2003

Dedication

This thesis is dedicated to my father and mother,

David and Margaret Williams, and my brother Mark.

Acknowledgements

I thank my advisor, Dr. Lloyd M. Davis for his continued assistance with all aspects of this thesis and the research undertaken herein. It has been the greatest pleasure to work under his guidance. I thank my other committee members, Dr. Christian G. Parigger and Dr. Horace W. Crater for their assistance in producing this thesis. I also thank my colleague David A. Ball, who has made many worthwhile contributions and collaborations to my research and other welcomed support. I express my gratitude to friends and family who have supported me through all my pursuits, with special thanks to Nancy Marshall and Sarah Kuhns for providing me with their extended hospitality during my time studying at UTSI.

Finally, I acknowledge financial support from the UTSI Center for Laser Applications, from LI-COR, Inc., Lincoln, Nebraska, with funding in turn from a phase II National Institutes of Health Nanotechnology SBIR award, from Abbott Laboratories, Chicago, Illinois, and from the UT Chemical Physics program for two travel grants.

Abstract

Computer modeling and analysis methods are developed for two modes of operation of an instrument for sensitive fluorescence detection of individual dye-labeled molecules in solution. First, Monte Carlo simulations of experiments for single-molecule imaging (SMI) are extended to include effects of sample flow, sticking of molecules to surfaces, and the finite depth-of-focus of the optics. The results have a bearing on a patented method for high-speed single-molecule DNA sequencing. They indicate that the imaging of freely moving fluorescent labels within a microfluidic flowcell will be considerably more involved than that of immobilized molecules at a surface, which is the usual situation in SMI experiments. Second, the detection of single molecules as they pass through a tightly focused laser beam is discussed, with an emphasis on fluorescence correlation spectroscopy and the analysis of the autocorrelation function of the photon counts. Analysis methods are developed and applied to data from a collaborative experimental study within the topic of RNA transcription. The methods are extended to the case of flowing solution, for ongoing research with application to high-throughput pharmaceutical drug screening.

Table of Contents

List of Tables

List of Figures

1. Introduction

1.1 Overview

Single-molecule detection (SMD) is a developing field of study, which has helped to forge a link between biology and physics. The first report of SMD experiments in the biosciences was in 1976 by Hirschfeld [1], where photon bursts were observed from fluorescent molecules attached to a protein. The technique was developed through the 1980s and 1990s [2,3,4,5,6,7] and has subsequently flourished into a widely used and studied discipline [8].

The method involves illuminating fluorescent molecules within a solution with a laser beam and collecting the emitted fluorescence with a detector such as a charged coupled device (CCD) or one or more single-photon avalanche diodes (SPADs). Different volumes of the solution may be illuminated dependent on the method used. For example, in the wide-field epi-illumination configuration, a broad beam is used to excite a planar sample of molecules, whereas in confocal epi-illumination, the laser beam is tightly focused so that volumes down to the sub-femtoliter scale may be probed. Total-internal-reflection (TIR) may also be used to excite the molecules, with the advantage of high signal-to-noise and reduced background [9].

The experimental apparatus for SMD at UTSI can be operated in the two above-mentioned configurations. The research conducted and presented in this thesis accordingly involves two subtopics within the field of SMD associated with different configurations. The subtopics are introduced in Sections 1.2 and 1.3 and elaborated in Chapters 2 and 3, respectively.

Section 1.2 discusses Single Molecule Imaging (SMI) methods, in which prism-coupled TIR is used to excite a solution of fluorescent molecules, and it presents images acquired in prior experiments with a CCD camera using this technique. The development of a computer simulation of these and similar experiments was a major component of this work. Such computer simulations are useful to aid the understanding of the experiments and the planning of future experiments [10]. Chapter 2 begins with a review of a SMI simulation initially created in MATLAB [11] by Dr. Greg R. Bashford and Dr. John G.K. Williams of LI-COR Inc., and subsequently augmented by Dr. Lloyd M. Davis and Dr. Kenneth R. Kimble of UTSI. Section 2.2 then presents subsequent refinements to the simulation, and the results of studies using the modified code, which contribute to the first part of the work for this thesis.

Section 1.3 introduces Fluorescence Correlation Spectroscopy (FCS) and related methods, which probe the kinetic properties of solution constituents by analyzing the sequence of photons individually detected from molecules as they diffuse through a tightly focused laser beam in the confocal epi-illumination configuration. An explanation is given of how FCS is used to determine the degree of binding of a fluorescently labeled ligand to a biomolecule in solution. The development of analysis methods for the autocorrelation function (ACF) of FCS data and the application of those methods to a collaborative biophysics experiment, and to ongoing research for high throughput pharmaceutical drug screening, contribute to the second part of the work for this thesis. Chapter 3 begins with a definition of the normalized ACF and presents a derivation of its theoretical form in the case of diffusional transport of molecules. The dependence of the shape of the theoretical form on the adjustable parameters is investigated in Section 3.2

and this is applied in Section 3.3 where curve-fitting is performed on a set of experimental data. A non-linear least squares curve fitting program, with a graphical user interface (GUI), was created in MATLAB specifically for determining the best fit parameters, and is presented in Section 3.4. In Section 3.5, the program is used to analyze data from experiments in collaboration between UTSI and UT Memphis to study protein binding within the process of DNA to RNA transcription. The theoretical form of the ACF is extended in Section 3.6 to consider the addition of a flow component to the molecule's diffusional motion and the dependence of the shape of the ACF on the flow velocity is examined in Section 3.7. Past experiments that incorporate flow are discussed in Section 3.8 along with the presentation of a Monte Carlo simulation [12] that models SMD experiments that include flow. Results produced by the simulation, in the form of the ACF, are fit and analyzed in Section 3.9.

Chapter 4 summarizes the conclusions from Chapters 2 and 3. Developments in the field during the course of this thesis are also briefly mentioned.

1.2 Single-Molecule Imaging

Ideas and methods of single-molecule imaging (SMI) date from at least 1988, when a system for tracking individual molecules within a flowcell was patented [3]. Since then much technological advancement, particularly in high sensitivity cameras, has been made, resulting in commercially available systems for SMI and many reports of single-molecule studies in the biosciences [8]. The development of SMI capability has also spurred research on several proposed methods for rapid DNA sequencing, in which

simultaneous SMD at many locations can be used to multiplex the sequencing process.

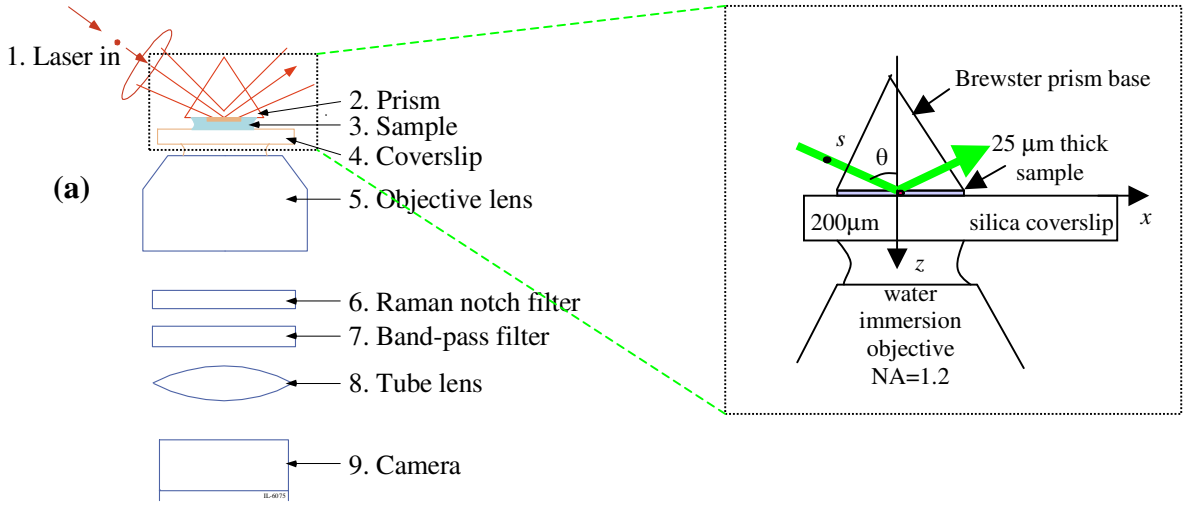
Whereas most SMI experiments require the molecules to be immobilized [13] or slowly moving [14], the rapid DNA sequencing method that has prompted this research [15] intends to use a flowcell environment, with molecules free to diffuse within it. A flow may be induced within the cell to cause molecules to pass more quickly through a specified region than by diffusion alone.

Several experimental studies have been made in support of SMI within a flowcell, an example of which is the one undertaken using the UTSI apparatus [16]. In one set of successful experiments the excitation laser beam is passed through a shutter, to yield a 3 ms pulse synchronized to the camera frame acquisition, in order to effectively take a snapshot of the diffusing molecules. Also, in order to minimize excitation of autofluorescence of optical components and achieve the highest signal-to-noise [9], the sample is illuminated via prism-coupled TIR, using a prism made of fused silica. Figure 1a illustrates the schematic experimental configuration. The refractive indices of fused silica ($n_1 \approx 1.46$) and water ($n_2 \approx 1.33$) at 568 nm yield a value for the critical angle, θ_C, into the silica-water boundary of

$$\theta_C = \arcsin\left(\frac{n_2}{n_1}\right) = 66.02°. \tag{1}$$

A beam with an angle of incidence $\theta > \theta_C$ results in TIR, and produces evanescent illumination within the flowcell, with exponentially decreasing irradiance with distance from the interface:

$$I(z) = I_0 e^{-z/d} \ , \tag{2}$$



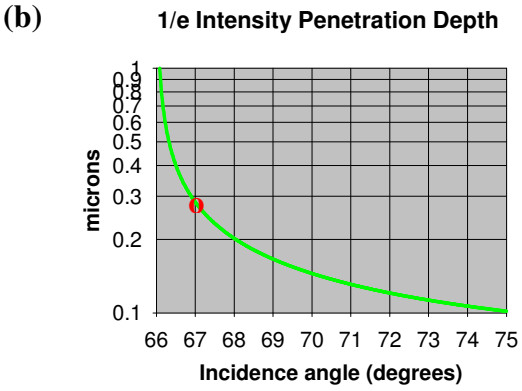


Figure 1: Prism-coupled total-internal-reflection excitation. (a) Schematic of set up illustrating geometry. (b) Evanescent field penetration depth versus incidence angle. (c) Evanescent field intensity enhancement factor versus incidence angle for s-polarized incident beam. Results presented in [16].

where I_0 is the irradiance at the interface within the sample, and d is the $1/e$ penetration depth. The dependence of d on θ is given by [16]

$$d = \frac{\lambda_0 / n_2}{4\pi\sqrt{\sin^2\theta - \sin^2\theta_C}}, \tag{3}$$

and is shown in Figure 1b. In order to increase the penetration depth, the angle of incidence should approach the critical angle. TIR of the incident laser beam at the surface also sets up a standing wave in the z-direction, which leads to an enhancement of the evanescent wave intensity. Figure 1c shows the θ dependence of the enhancement factor, which is given by

$$I_0 / I_s = \frac{4\cos^2\theta}{1-(n_2/n_1)^2}, \tag{4}$$

where I_s is the irradiance of the s-polarized incident laser beam. For the experiments reported in reference [16], the enhancement factor is about 3.7, and the penetration depth is $d = 0.3$ μm.

Experiments were conducted for many samples and sample concentrations over a range of laser powers. One typical example, as given in [16], studies a sample of unconjugated Bodipy-Texas Red dye (Molecular Probes, Inc.), at a concentration of 2.8×10^{-11} M in water. The sample is illuminated via TIR using 568 nm Ar/Kr ion laser radiation. Figure 2 shows a series of images taken with an intensified CCD (ICCD) camera (Roper Scientific Pentamax). The images show that molecules generally come and go as they diffuse in and out of the evanescent field (see downward-pointing arrows), with others suddenly appearing and then either remaining stuck to the interface (see dashed arrows, frames 2.8s – 3.4s), or remaining stuck for a shorter period of time (see open left-pointing

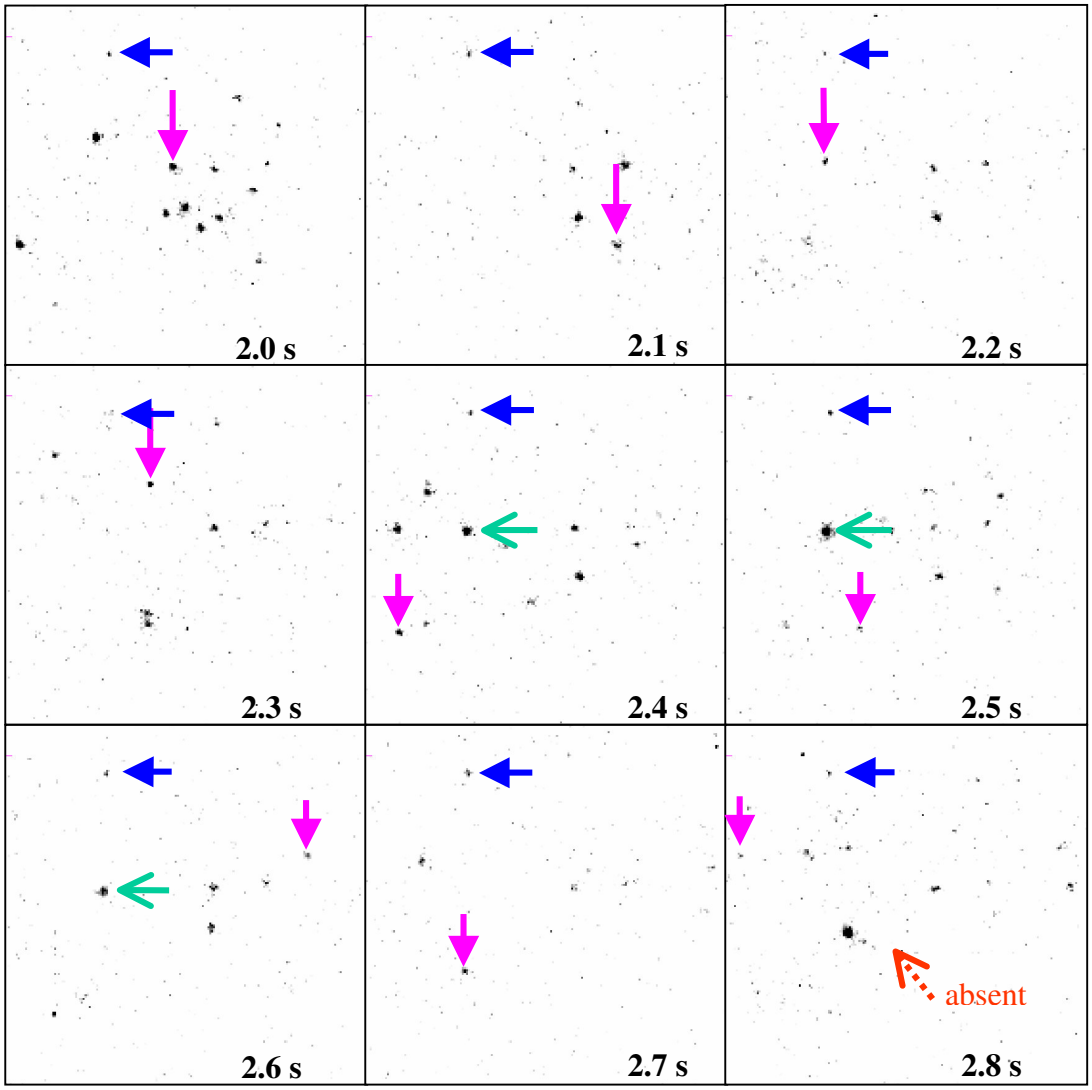


Figure 2: Sequence of images obtained from 3ms laser exposure of Bodipy-Texas Red solution. Each frame is separated by 100 ms with an image magnification of 157.5. Reproduced from [16].

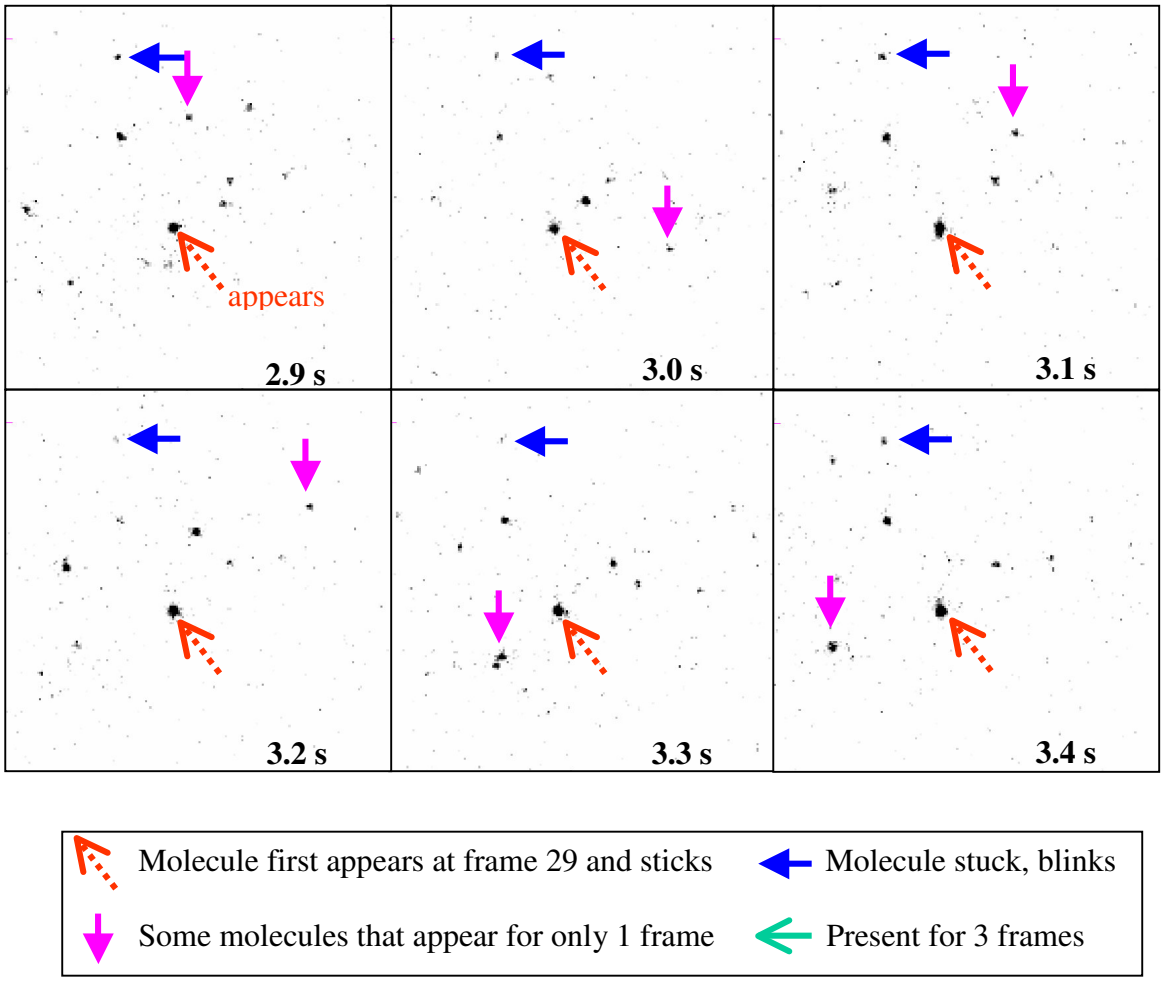


Figure 2: Continued.

arrows, frames 2.4s – 2.6s) before becoming photo-bleached [17], and thus unobservable. Molecules that are stuck to the interface also show signs of blinking (see solid left-pointing arrows).

Chapter 2 reports Monte Carlo simulations of experiments similar to those of reference [16], including refinements to the simulation to model the molecule sticking behavior seen in those experiments, and other refinements needed to determine anticipated results within a flowcell.

1.3 Fluorescence Correlation Spectroscopy

FCS is a resourceful method of determining the components of a solution and their kinetic and photophysical properties from measurements of the fluorescence fluctuations that arise when concentrations are low, i.e., observed molecule numbers are at the single-molecule level. Since its early inception [18], recent technological developments in the area of SMD have enabled FCS to become a practical tool in the biosciences, with second-generation instruments now commercially available [19]. In these and in the UTSI instrument for SMD [20], a confocal epi-illumination configuration is used to define a sub-femtoliter probe volume so that enhanced signal-to-noise data can be collected. The laser excites a molecule diffusing into this illuminated probe volume. On de-excitation, the molecule emits a fluorescence photon, which may be collected by a SPAD detector and accumulated as a single photon count. During passage through the probe volume, a single molecule may undergo thousands of excitations and photon emissions, resulting in a burst of photons. During the course of an experiment, many

fluorescent molecules pass through the probe volume, culminating in a series of photon bursts.

There are various methods of analyzing the photon burst data, which may yield various experimental parameters. In the usual data analysis method [5], each detected photon is correlated with every preceding one by accumulation of a histogram of delay times between the photons to form the autocorrelation function (ACF) of the stream of photon counts. Information on the time scales of the fluorescence fluctuations and on the concentrations of solution constituents that yield the different time-scale fluctuations may be derived from the ACF.

Other data analysis methods are also possible. In Successive Photon Interval Density Analysis (SPIDA) [21], one builds up a family of histograms of time intervals between increasing numbers of successive photons. If all the histograms are added together, one obtains the ACF. In Fluorescence Intensity Distribution Analysis (FIDA) [22], one collects a histogram of the number of photons within successive time bins of fixed width. This histogram gives information on the fluorescence brightness values of solution constituents. In Fluorescence Intensity Multiple Distributions Analysis (FIMDA), the FIDA method is used, but a family of histograms is accumulated for a range of temporal bin-widths [23]. Thus, in FIMDA, one captures information on both the times scales and the brightness values of the fluorescence fluctuations.

The ACF, or one of the other data analysis methods, may be used to quantify interactions between biomolecules as follows: A solution contains a target biomolecule to which a fluorescently labeled ligand is bound. FCS is used to determine the kinetic properties of this molecular pair. When a drug-like compound to be tested is added to the solution, the

FCS data would be the same as before if no binding were to take place. However, if binding does occur, the ligand becomes displayed from the biomolecule. In the extreme event of 100% binding, all fluorescently labeled ligands are now free and an FCS measurement would yield different data, as the brightness or rate of diffusion of the ligand are generally different when the ligand is unattached. In the case of the ACF, the main kinetic characteristic to change is the diffusion constant, D, of the ligand. For Einstein-Stokes diffusion, D is approximately inversely proportional to the cube root of the molecular weight. Hence if binding occurs, the ACF would indicate shorter photons bursts due to the faster diffusion of the unbound ligand.

Chapter 3 presents a derivation of a theoretical model of the ACF, which is then used to derive kinetic properties of the diffusing molecules.

2. Simulations of Single-Molecule Imaging

2.1. Initial Simulation Code

Section 1.2 introduced experimental methods for the imaging of single molecules and presented work carried out by Davis *et al*. [16] in which single molecules are visible in a series of images. It is natural that one would want to refine the experiments to gain better results, such as improved images, or to adopt the experiment to one's needs, for example, by the addition of flow. As an aid to doing this, a Monte Carlo simulation program was written using MATLAB by G.R. Bashford and J.G.K. Williams of LI-COR, and then expanded by L.M. Davis and K.R. Kimble of UTSI to include TIR illumination, to simulate freely diffusing molecules, and to include the light detection and read out characteristics of a back-illuminated CCD camera. This expanded MATLAB simulation, which gave results in qualitative agreement with the experiments in [16], formed the starting point for the research reported in this chapter. The program code listing is given in Appendix 1. Subsequent updates are labeled and referenced by numbers in braces (e.g., {1}) when each is discussed later in the chapter.

The SMI experiment under simulation in the initial program is for molecules of Bodipy-Texas Red dye at a concentration of 6.15×10^{-10} M, free to diffuse in water with $D = 2.8\times10^{-6}$ cm^2 s^{-1}. Molecules are confined to a simulation volume 25 μm deep in the z-direction. The volume is also restricted in x and y by the number of CCD pixels, which is set at 128 for both x and y. The size of the volume is then dependent on the pixel size of the camera and the optical magnification. An optical magnification of 157.5 was used. Should an ICCD be used (as investigated in [16]), the pixel size would be 22.5 μm, and

the x and y lengths would be 18.3 μm. The simulations in this chapter, however, investigate a back-illuminated CCD with a pixel size of 13 μm, which results in x and y lengths of 10.5 μm. If a diffusing molecule impinges on a boundary, it will simply bounce, without sticking to the interface.

The simulation also models the TIR illumination of the sample with an incident laser beam with a wavelength of 568 nm and a power of 325 mW, which is pulsed once every 280 ms for a duration of 3 ms, as explained in the description of the experiment in Section 1.2. The penetration depth and the enhancement factor of the evanescent wave are 0.3 μm and 3.7, respectively. As seen in the code listing in Appendix 1, various other optical parameters are included. Most notable of these is that the focal plane of the objective lens through which fluorescence is collected is set at the substrate surface, i.e., at z = 0, and the fluorescence from a point source such as a molecule undergoes diffraction due to the limited numerical aperture of the objective. The resultant Airy disk is approximated by a two-dimension Gaussian with a spot radius in object space given by

$$\sigma = \frac{0.4207\lambda}{2NA}, \tag{5}$$

where λ is the emission wavelength, NA is the numerical aperture of the objective lens, and the factor of 0.4207 arises due to the Gaussian approximation of the Airy function. The camera pixel that receives the fluorescence photon from a given molecule is determined by the x,y-location of the molecule in the simulation volume, plus a Gaussian distributed random number with a standard deviation given by equation (5). With λ = 568 nm and NA = 1.2, the spot radius in object space, i.e., the standard deviation found from equation (5), is 0.107 μm, which corresponds to 1.3 pixels in image space.

When the program is executed, the simulation goes through the iterations shown in the flow diagram of Figure 3. The simulation results in a series of images, the first six of which are shown in Figure 4. The colored bar to the right of each frame shows an arbitrary fluorescence intensity scale, with white showing little or no fluorescence and black showing areas of high fluorescence. Throughout the series, dark patches that correspond to the detected fluorescence from individual molecules can clearly be seen.

2.2. Simulation Modifications and Investigations

2.2.1. Increased Exposure Time for a Pulsed Laser

The laser exposure time of 3 ms initially used in the simulation to model the experiments described in Section 1.2 effectively takes a snapshot of the rapidly diffusing molecules. In order to study what would be seen in experiments with a back-illuminated CCD with exposure times increased beyond 3 ms, simulations were carried out for a range of exposure times up to the point of continuous exposure. It is expected that smearing effects would begin to take place, analogous to opening a camera shutter for a lengthened period, and would eventually degrade the molecule images.

Due to the long run time of the code, the simulations were executed to capture only two CCD camera frames, by use of the following sequence:

Iterations with laser on (capture CCD frame 1);

Iterations with laser off;

Iterations with laser on (capture CCD frame 2).

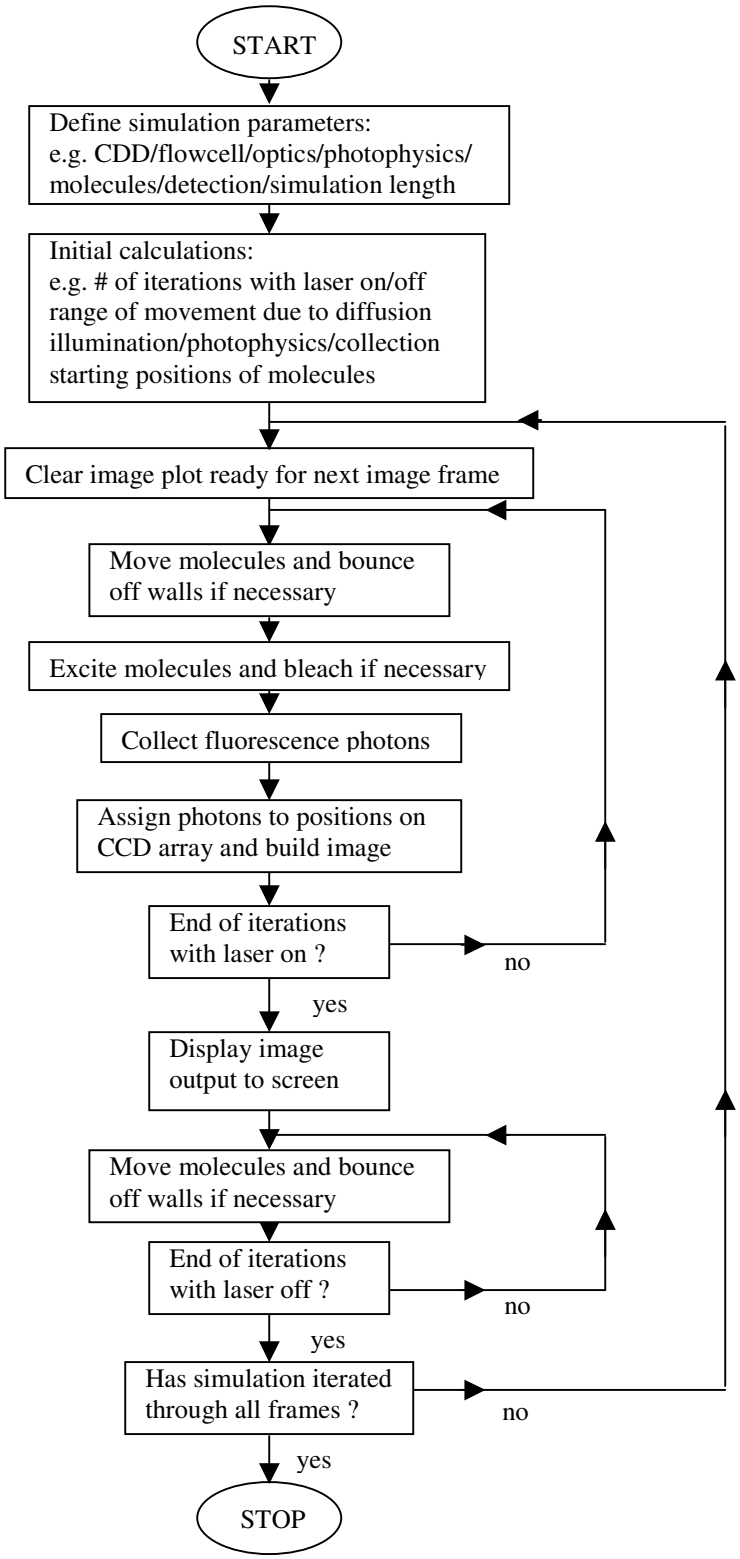
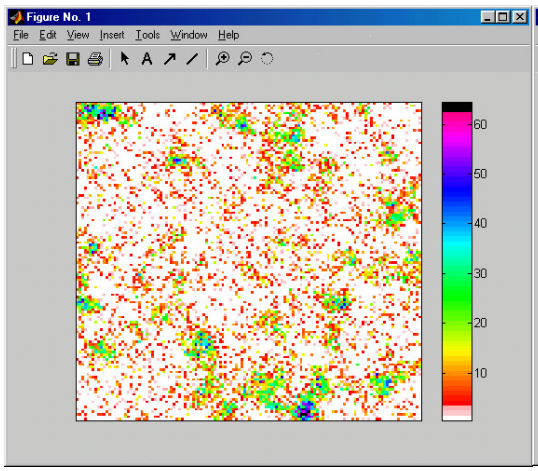
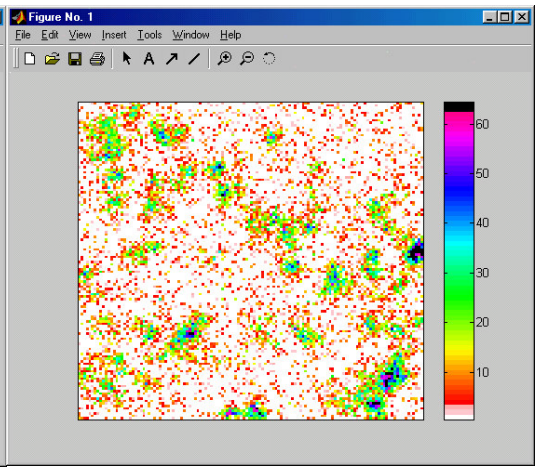


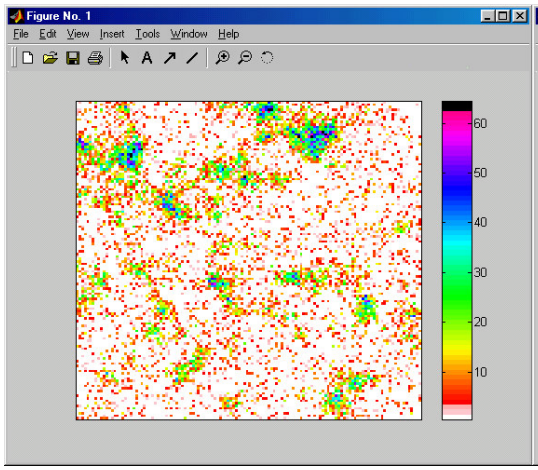
Figure 3: Flowchart for a Monte Carlo simulation of single-molecule imaging.



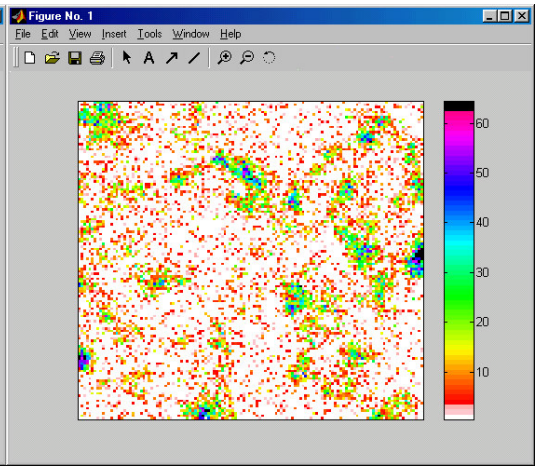
Frame 1 (0.28 s)



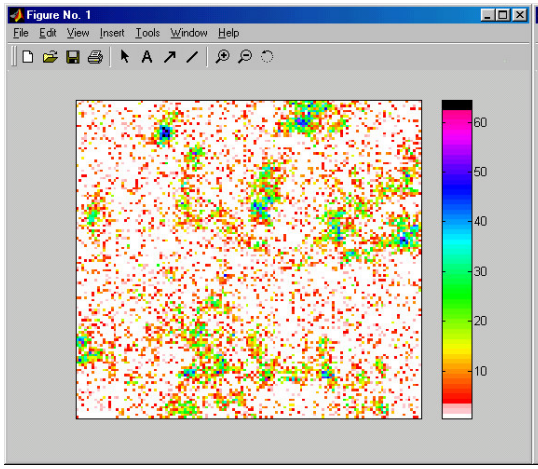
Frame 2 (0.56 s)



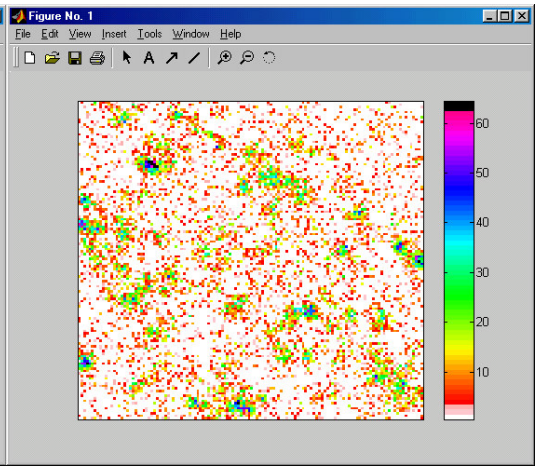
Frame 3 (0.84 s)



Frame 4 (1.12 s)



Frame 5 (1.40 s)



Frame 6 (1.68 s)

Figure 4: Series of image frames produced by MATLAB Monte Carlo simulation.

Table 1 lists the laser exposure time (LET), the number of iterations carried out with the laser on, prior to capture of the first frame, the number of iterations with the laser off, the number of molecules in that particular simulation, Nm, a Poissonian random number proportional to the flowcell dimensions and the molarity (given in Section 2.1), and the execution time on a 766 MHz Pentium 3 PC. For continuous laser exposure, in which the photophysics must be modeled in all numerical iterations, the slow simulation execution time becomes an issue for any investigations.

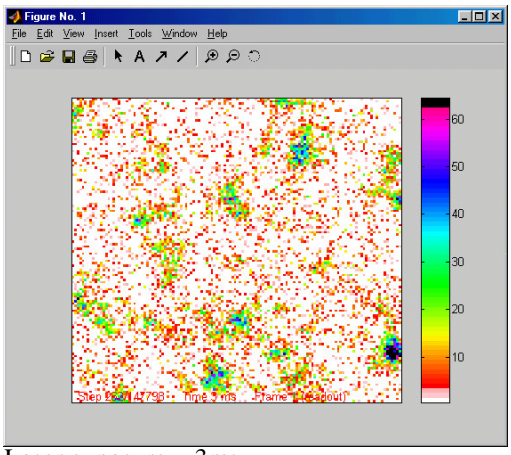
The frames for each simulation are shown in Figure 5. It can be seen that as the laser exposure is increased, the fluorescence signal increases, but smearing of the image also increases. For continuous exposure, one frame capture is adequate to show that the image becomes almost saturated and that applications to the experiment at this laser power level are not useful.

2.2.2. Continuous Laser Exposure with Varied Laser Power

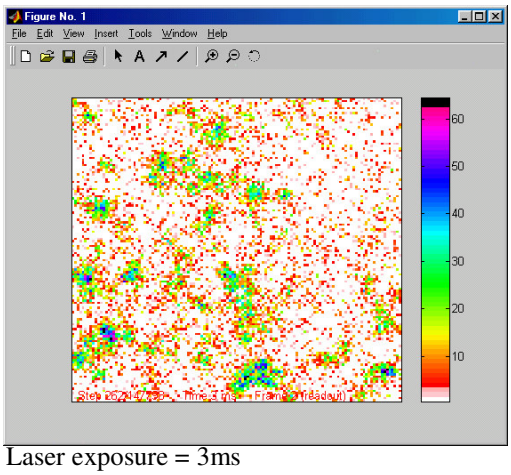
Experiments conducted on the UTSI SMI instrument [16, 20] and on an instrument at LICOR, Inc., had demonstrated that clear images of single molecules could be obtained for continuous laser illumination with milliwatt power levels. However, the extent to which freely diffusing molecules contributed to these images was unclear. Therefore, simulation investigations were made to determine what would be expected for continuous laser exposure of freely diffusing molecules, but with powers below the initial level of 325 mW. With reduced power, the intensity of the images is expected to decrease, which acts to counter the high intensity fluorescence due to continuous laser exposure.

Table 1: The dependence of the execution time on the laser exposure time (LET).

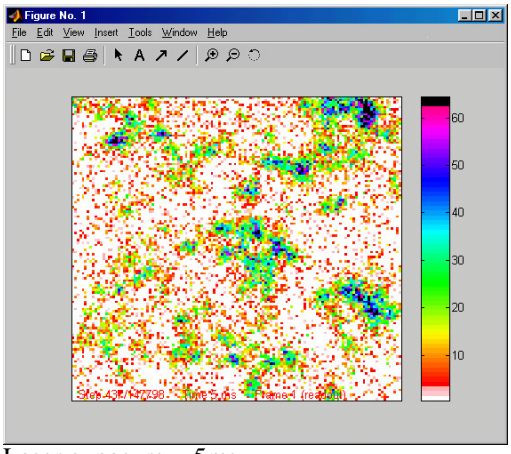
LET (ms)	Iterations w/ laser	Iterations w/o laser	Nm	Time (mins)
3	262	24370	1016	3
5	437	24195	1046	4
10	875	23758	1026	5
20	1750	22883	1025	9
50	4375	20258	1003	20
280	24500	133	1032	95



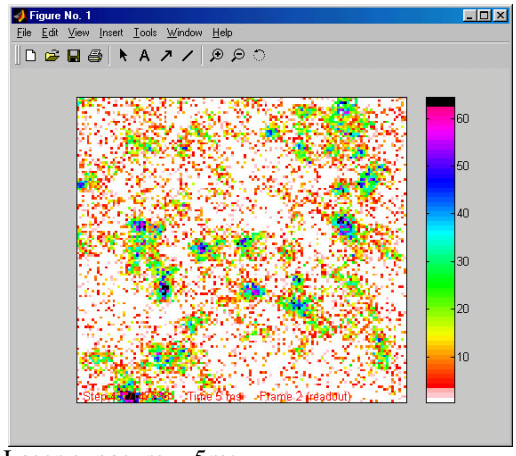
Laser exposure = 3ms
Frame 1



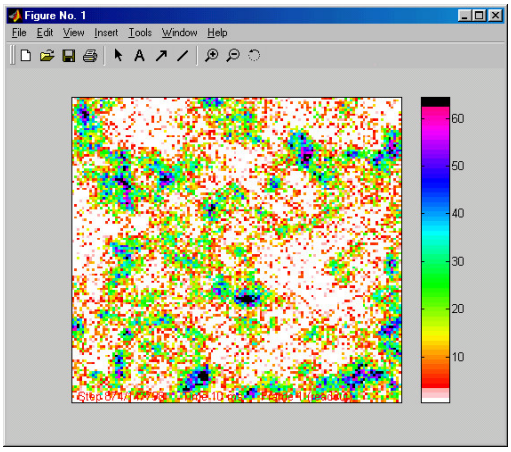
Laser exposure = 3ms
Frame 2



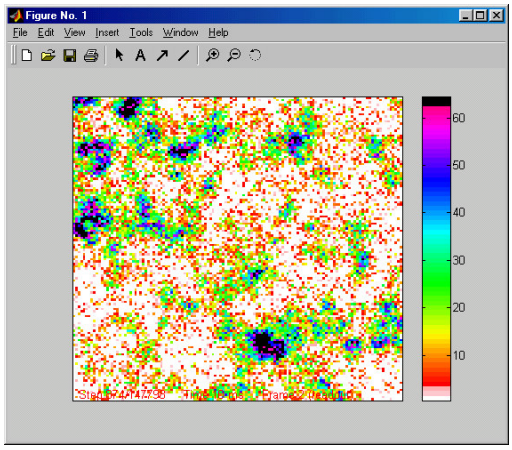
Laser exposure = 5ms
Frame 1



Laser exposure = 5ms
Frame 2

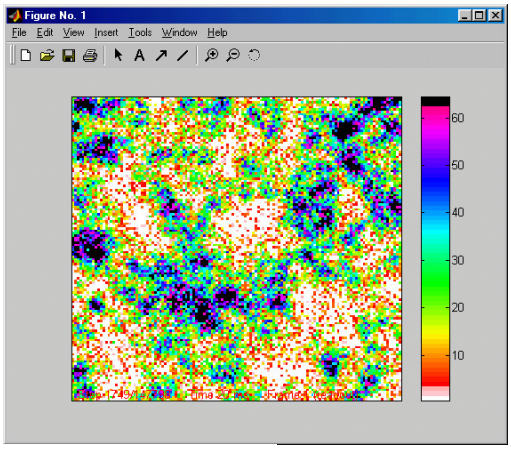


Laser exposure = 10ms
Frame 1

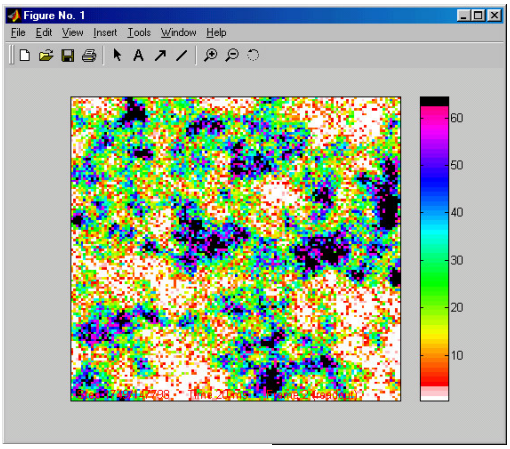


Laser exposure = 10ms
Frame 2

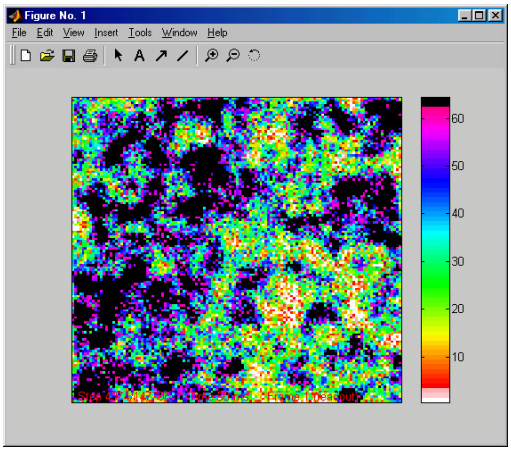
Figure 5: Effect of exposure time on image.



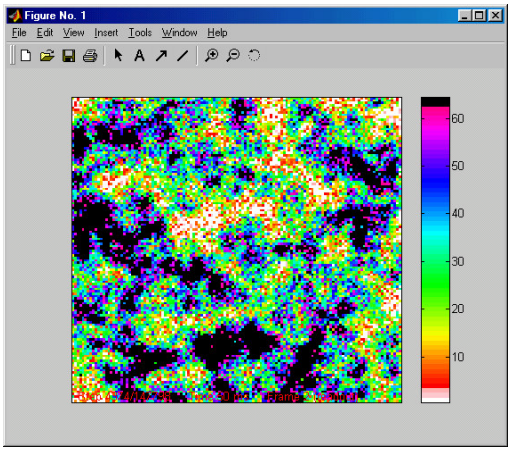
Laser exposure = 20ms
Frame 1



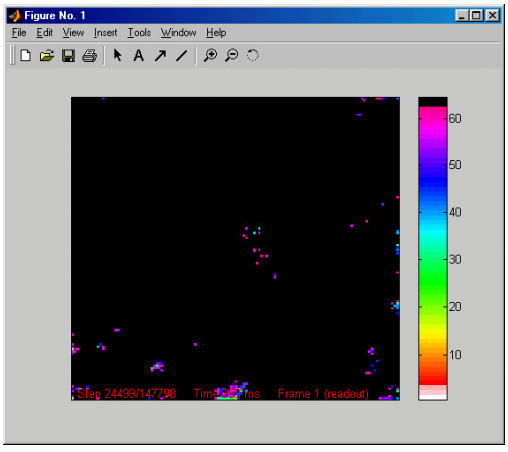
Laser exposure = 20ms
Frame 2



Laser exposure = 50ms
Frame 1



Laser exposure = 50ms
Frame 2



Laser exposure = 280ms
Frame 1

Figure 5: Continued.

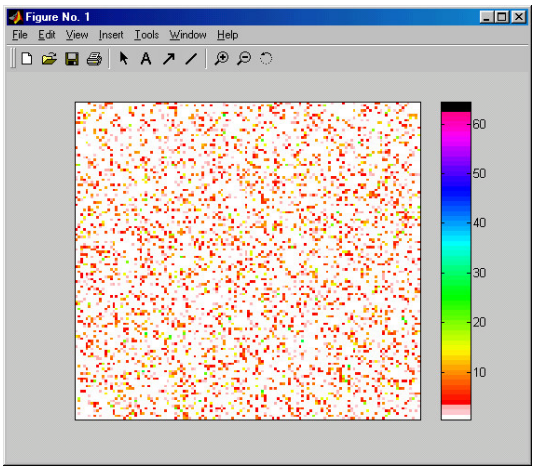
A series of simulations was run with constant laser exposure at different laser powers. The first frame captured of each respective simulation is shown as a set of images in Figure 6. The laser power was initially at the μW level, and was then increased for subsequent simulations. It was found that nothing interesting was imaged for laser powers of below ~ 1 mW. Increased powers show structure with evidence of smearing (see 10 mW image in Figure 6). At powers of 50 mW the images still exhibit smearing although are more intense. With increased powers, results (not shown here) are similar to that of the final exposure of Figure 5, presented in the previous Section. The observations show that although continuous exposure for a greatly reduced laser power still produces images containing fluorescence, no discernable molecule images are visible due to smearing effects.

2.2.3. Investigation of Molecules Sticking to the Substrate

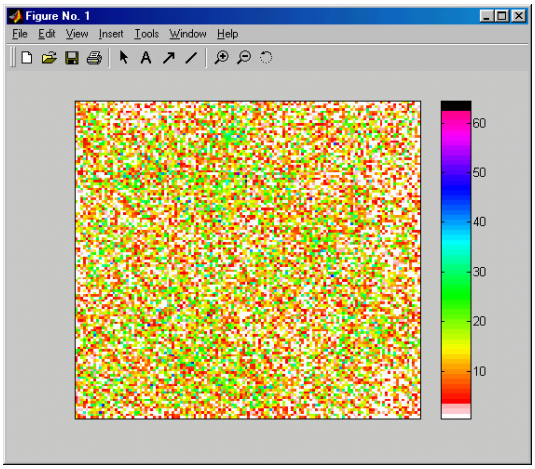
2.2.3.1 Implementation of Sticking in the Simulation

Evidence of molecules sticking to a flowcell substrate has already been observed [16, 20] and thus provides motivation for the implementation of such characteristics into the present simulation.

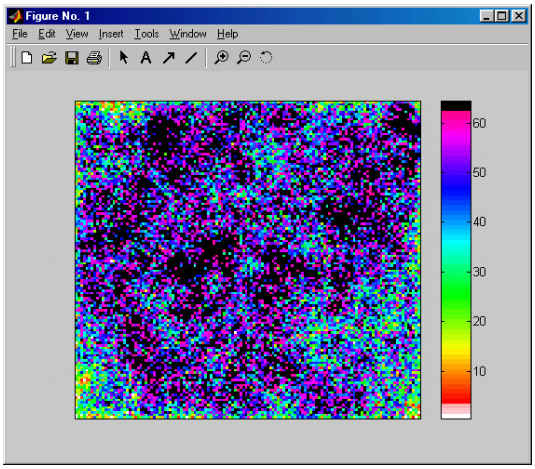
Molecules that diffuse to the substrate boundary (the $z = 0$ flowcell wall), which simply bounce in the simulations presented in previous sections, are now subject to a random number determination as to whether they will bounce or stick. Molecules that would have crossed the z=0 boundary are found, and those that are randomly determined to not bounce have their z-coordinate set to zero for the remainder of the simulation {1}.



Number of molecules = 1016
Laser power = 1 mW
Laser intensity = 217.0822 W cm^{-2}



Number of molecules = 1028
Laser power = 10 mW
Laser intensity = 2170.8 W cm^{-2}



Number of molecules = 1098
Laser power = 50 mW
Laser intensity = 10854 W cm^{-2}

Figure 6: Images to illustrate the effect of increased laser power.

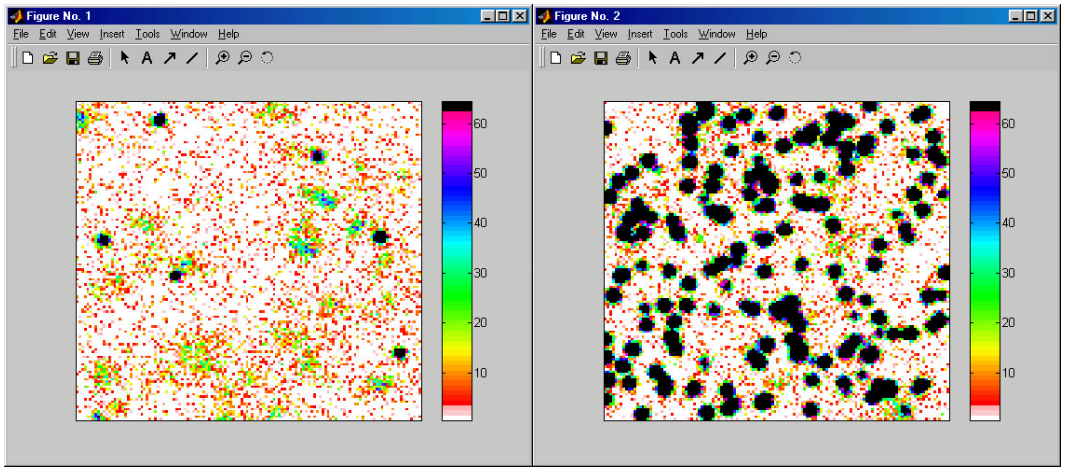
The probability of sticking is an adjustable parameter initialized at the start of the simulation {3}.

Once a molecule is stuck to the substrate, it must also be stopped from moving in the *x* and *y* dimensions. Thus the code in which the molecule positions are updated must also be altered to move only those not stuck to the substrate {6}, with similar coding for *y* and *z*. These changes are made within the iterations for which the laser is both on and off.

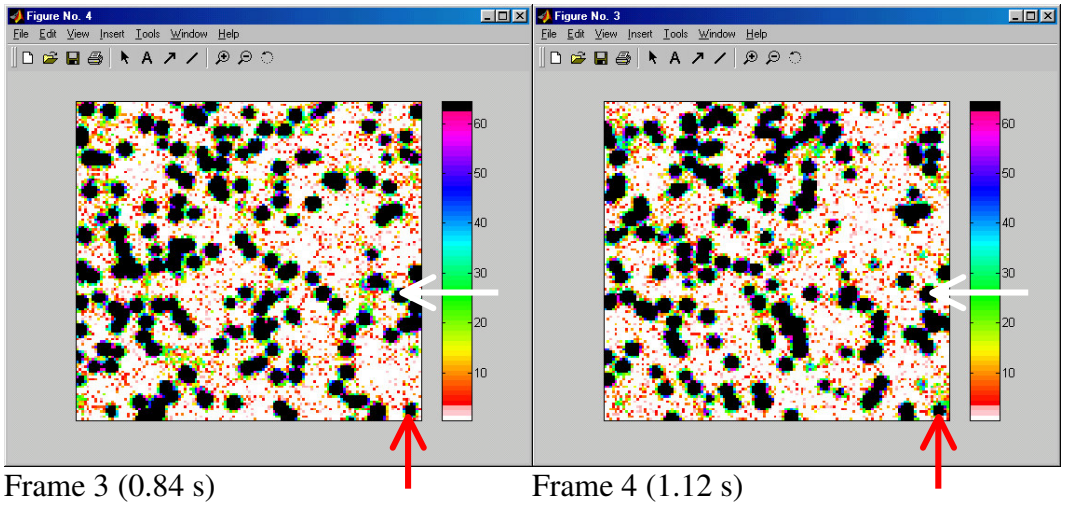
Note that molecules that stick or photo-bleach are not replaced in the simulation, and thus the concentration in free solution decreases, as would occur in a flow cell with small finite volume.

2.2.3.2 Results for Pulsed Laser Exposure

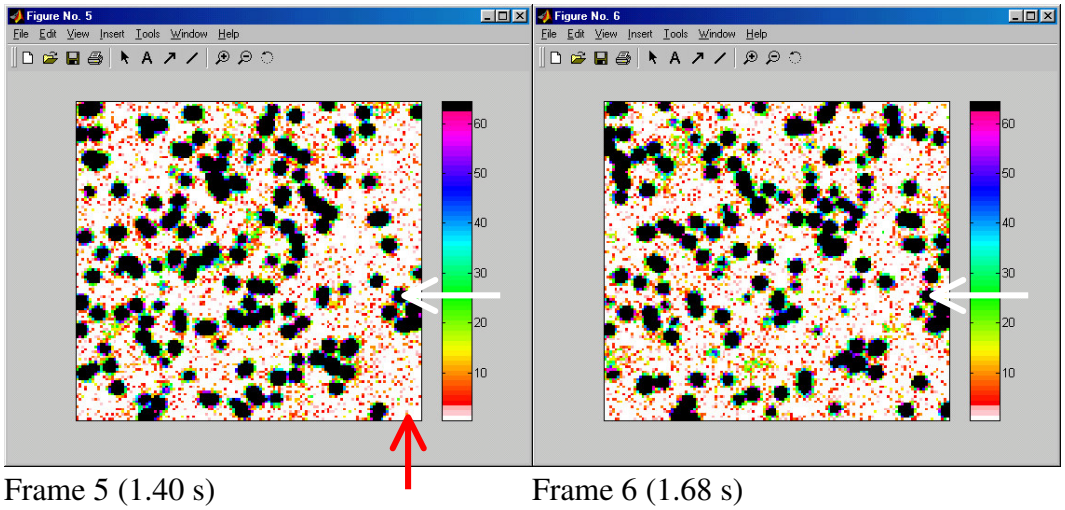
Figure 7 shows a series of six images with sticking at 10%. Although some molecules are seen to stick and remain visible throughout the simulation (left-pointing arrow), others are seen to stick and then disappear from the image (up-pointing arrow) after they are photo-bleached. For comparison, the images shown in Figure 4 were generated with no molecule sticking implemented in the code. Figure 8 shows the first two frames of three sticking probabilities, 0%, 50%, and 100%. Again, the 0% situation is similar to that shown in Figure 4 and acts as a control. The 50% case shows some molecules stick even after one frame, with the number increasing in the second frame. The 100% case looks similar to that for 50%. Although one might expect more molecules to be stuck for the 100% case, there are several reasons why this does not happen. First, if a molecule does not stick when it initially crosses the boundary, it has almost 50% probability to diffuse



Frame 1 (0.28 s) Frame 2 (0.56 s)

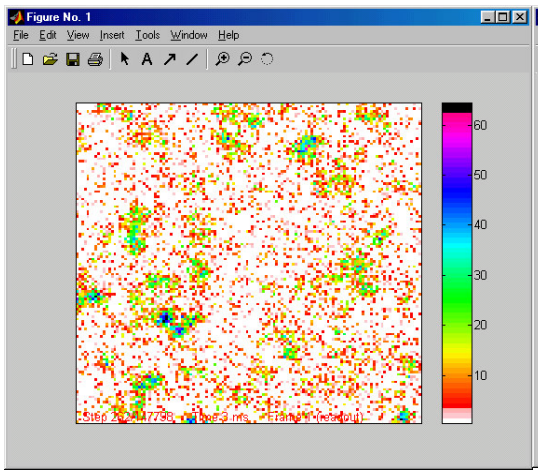


Frame 3 (0.84 s) Frame 4 (1.12 s)

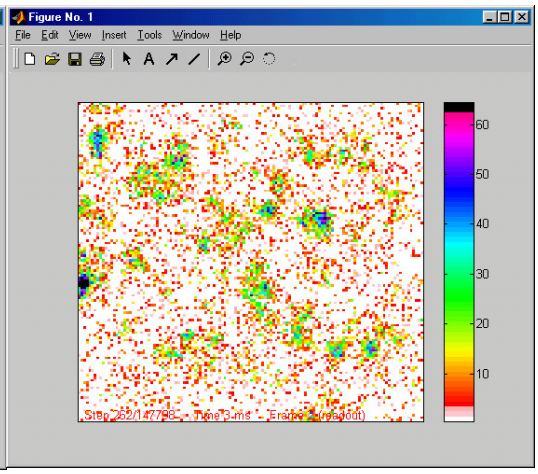


Frame 5 (1.40 s) Frame 6 (1.68 s)

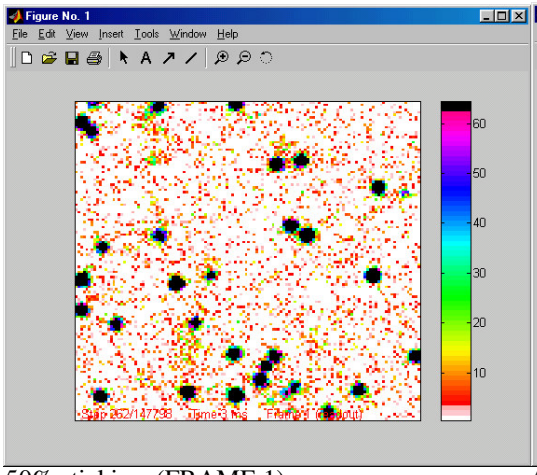
Figure 7: Series of images to show molecules sticking to the substrate.



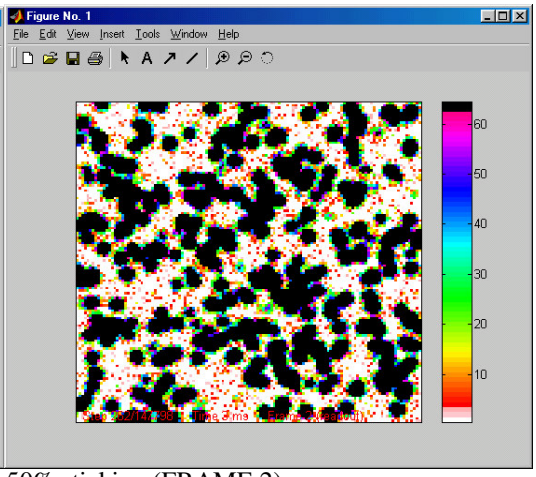
0% sticking (FRAME 1)



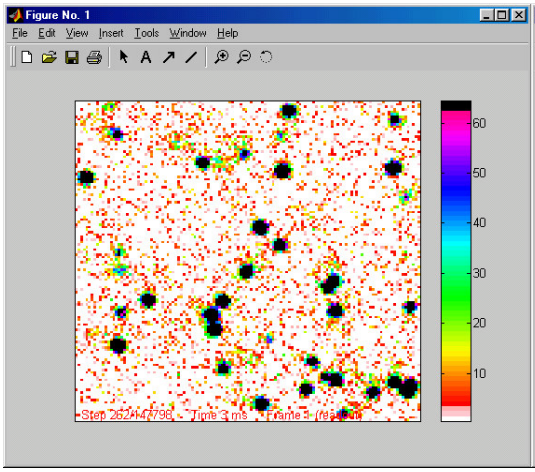
0% sticking (FRAME 2)



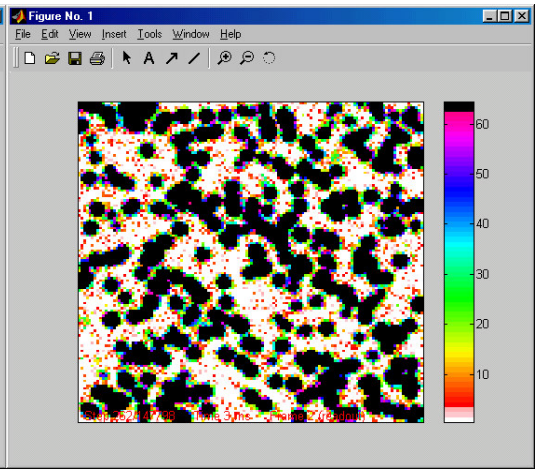
50% sticking (FRAME 1)



50% sticking (FRAME 2)



100% sticking (FRAME 1)



100% sticking (FRAME 2)

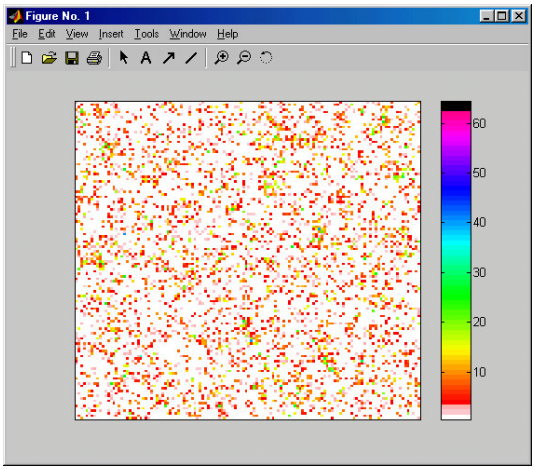
Figure 8: First two frame images for different sticking probabilities.

back to the boundary in the next iteration, thereby compounding the chances of sticking. Second, the number of molecules close to the boundary is finite. Hence if most of these become stuck, a higher sticking probability will not substantially increase the number that stick. Third, photo-bleaching of molecules that stick reduces the number visible in subsequent frames, although this effect is not as significant as the first two reasons.

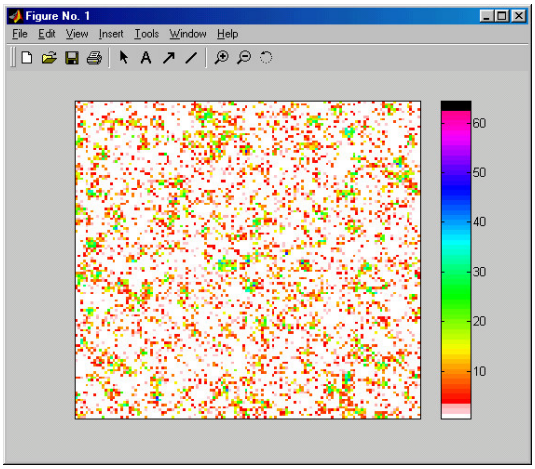
2.2.3.3 Results for Continuous Laser Exposure

The simulations in Section 2.2.2 show that fluorescence was seen under continuous laser exposure but the images of the molecules themselves were smeared and not discernable. Molecules stuck to the substrate would not move during the simulation and thus should be visible when continuously exposed by the laser. It is useful to find out whether this is indeed the case and at what laser power these stuck molecules would be seen.

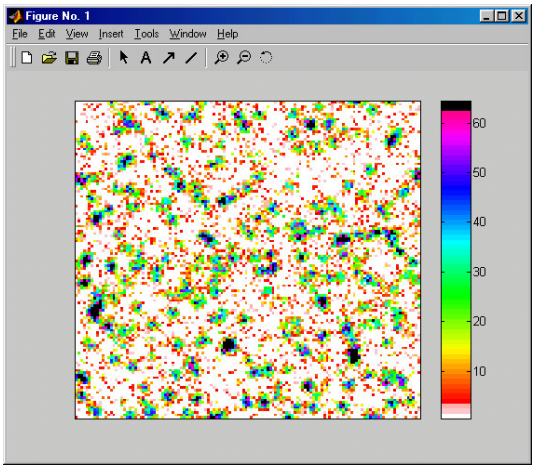
For this set of simulations, the sticking probability is set to 100%. The images referred to below, which are displayed in Figure 9, capture the first frame in each respective simulation. As before, the laser power was decreased into the μW region and then increased to 50 μW, where very faint structure may be seen. This structure is more discernable at a power of 100 μW. On comparison with the simulations without sticking described in Section 2.2.2, where no molecules were seen at 100 μW, the present results indicate that the only molecules that may be imaged at μW powers are those stuck to the substrate surface. Further enhancement of the molecule images is seen for further increase in laser power up to a level of 1 mW, where the images become similar to those seen in Figure 7.



Number of molecules = 1005
Laser power = 50 µW
Laser intensity = 10.8541 W cm^{-2}



Number of molecules = 1033
Laser power = 100 µW
Laser intensity = 21.7082 W cm^{-2}



Number of molecules = 1027
Laser power = 300 µW
Laser intensity = 65.1247 W cm^{-2}

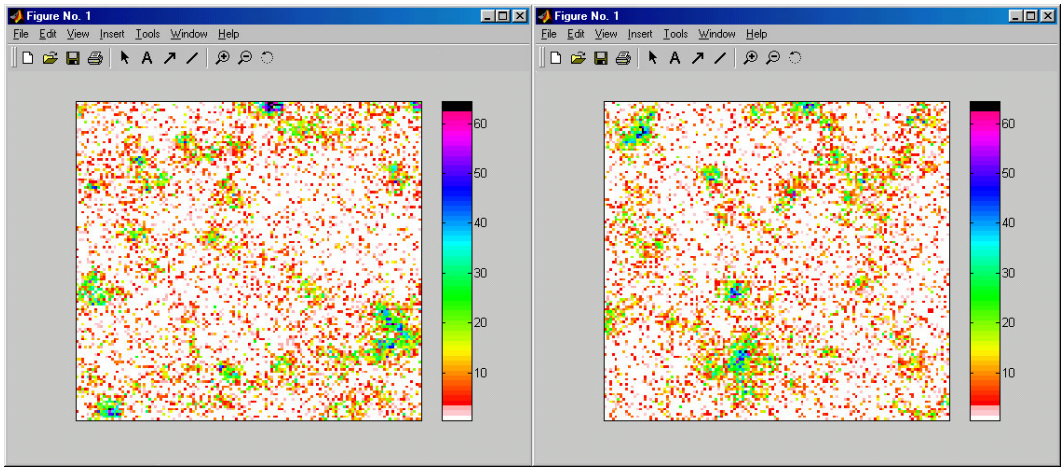
Figure 9: Effect of increasing laser power from low levels to those comparable to the initial simulations with molecules stuck on the surface.

2.2.4. Defocus Effects

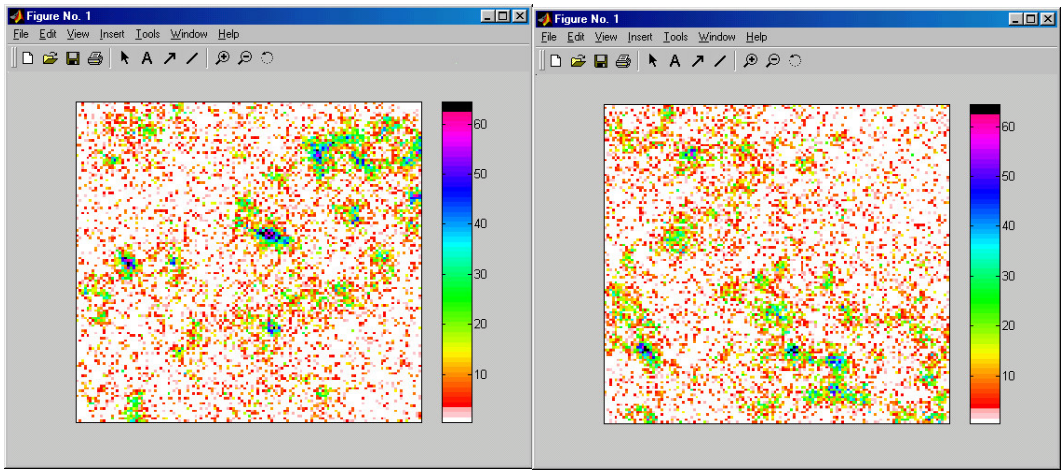
In the simulations presented so far, the plane of focus for the objective lens was set at the substrate surface, i.e., $z = 0$, and any effects of defocus of images of the molecules away from that plane were neglected. To further improve the simulation, such defocus effects are included. These effects were modeled using Zemax optical design software as a part of this thesis research. The results show that a defocus in either direction away from the focal plane yields a disk image with equivalent radius in object space of ~2.07 μm per μm defocus {2}. The distance from the focal plane at which defocus becomes important, which is found by comparison of the disk image radius with the 0.1069 μm object-space radius of the Airy disk, is found to be 0.0516 μm {4}. Thus, any molecules located at a z-distance from the focal plane greater than this value will be subject to image defocus and will not be diffraction limited. Note that 0.0516 μm is much less than the 0.3 μm penetration depth of the evanescent wave, and hence defocus effects should be considered even when using TIR-illumination.

Figure 10 shows a series of frames from a simulation with these defocus effects implemented. On comparison with the initial simulations (Figure 4), it may be seen that the molecule image structure is indeed more spread out, as would be expected from defocus effects.

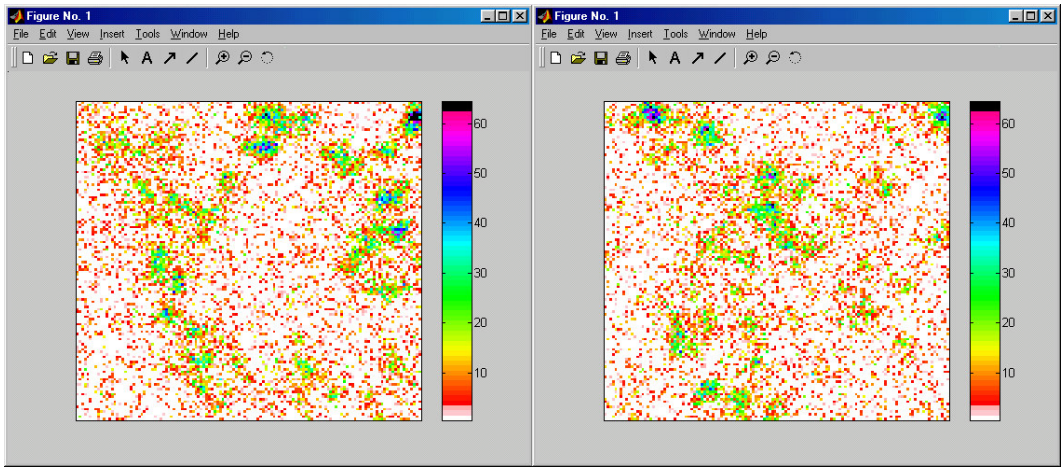
With molecule sticking implemented, the difference between the Airy and defocus disk image sizes are clear. In Figure 11, the first image frame of the simulation is shown and it is apparent that the more intense images of the stuck molecules are smaller than those of the free molecules that lay outside the defocus threshold.



FRAME 1 (0.28 s) FRAME 2 (0.56 s)



FRAME 3 (0.84 s) FRAME 4 (1.12 s)



FRAME 5 (1.40 s) FRAME 6 (1.68 s)

Figure 10: Effects of defocus on molecule images.

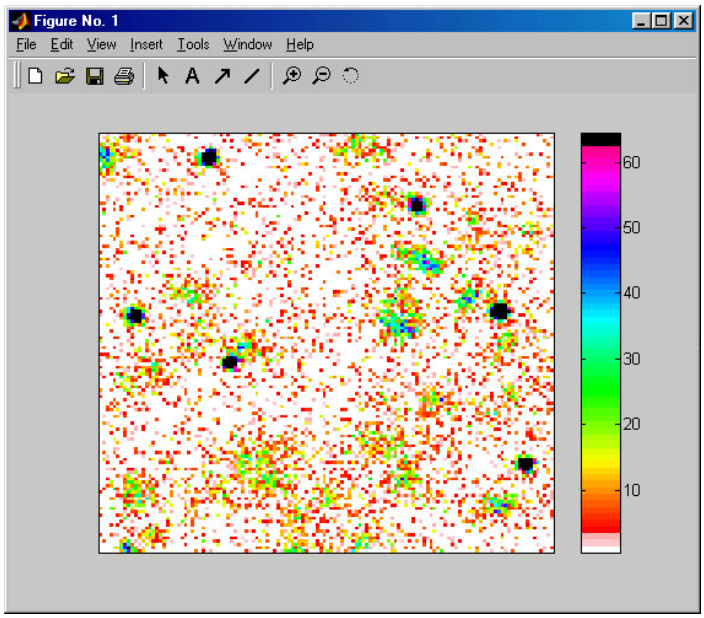


Figure 11: Effects of defocus on non-stuck molecule images. The image shows the distinction between the sizes of those images and those of the stuck molecules, which only undergo Airy diffraction.

2.2.5. Wide-Field Epi-Illumination of a Thin Flowcell

For DNA sequencing experiments, fluorescent molecules would be confined to a thin flowcell and excited via epi-illumination, so that the intensity of radiation is constant throughout the cell. This could be implemented in the code by use of an infinite penetration depth, so that equation (1) simply becomes $I(z) = I_0$. However, to speed the implementation of epi-illumination, the value of the mean probability of excitation per time-step is set to be a constant independent of z for all molecules {9}. The depth of the flowcell is also reduced from 25.0 to 2.0 μm and the TIR enhancement factor is now set to 1.

In all previous simulations, the objective lens has been focused at the $z = 0$ plane, i.e., the substrate interface. However, in order to more clearly image molecules within the flowcell, it is expected that the focal plane should be repositioned to the center of the flowcell. Defocus effects are expected to increasingly degrade the molecule images the further they are located from the focal plane, and the simulation is used to investigate the extent of this degradation.

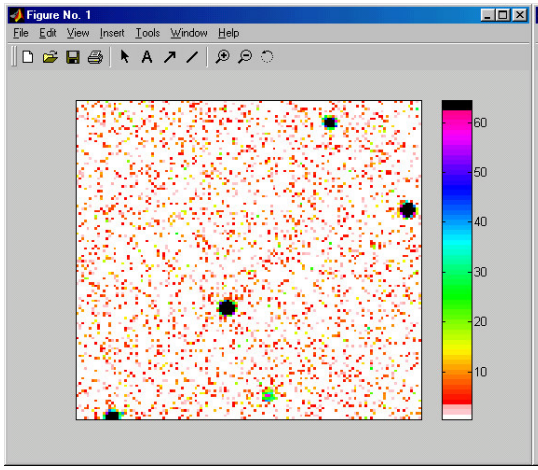
A parameter '*zfocus*' was introduced to set the z-position of the focal plane. Any molecule within a distance of 0.0516 μm from *zfocus* will be diffraction limited {10}, whereas molecules laying further away will have their images defocused {11}, where the radius of the defocused disk depends on the distance between the molecule and the focal plane {12}. The photons from the Airy disk or defocused disk are then collected {13} and displayed to screen.

Figure 12 shows a series of images of stuck molecules over a range of *zfocus* values. The sticking probability was set to 100% to see how the disk of the stuck molecule image varies as *zfocus* is increased toward the center of the flowcell. Each image captures the first frame in each respective simulation, which used the parameter values given in Section 2.1, including 3 ms pulsed excitation, apart from these changes:

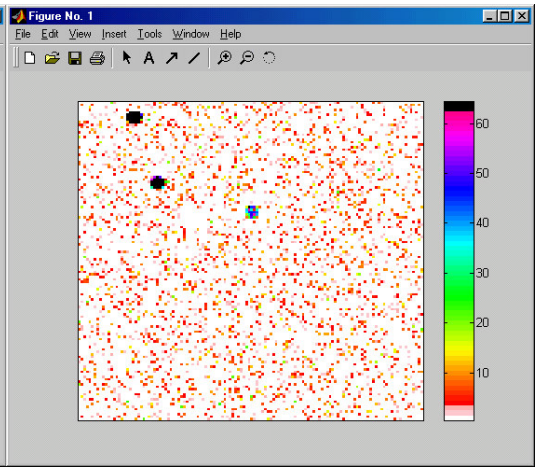
- Molarity, M = 6.15×10^{-11} M
- Sticking = 100% (note that sticking only occurs on the bottom interface)
- Flowcell depth = 2.0 μm

As *zfocus* is increased, the defocused image radius increases as expected. At the center of the flowcell, *zfocus* = 1.0 μm, it was found that the images of the molecules stuck to the bottom interface become so spread out that they are not seen. This may be considered to be an advantage, as they would not obscure the images of molecules at the center of the flowcell, although they would contribute to background luminescence.

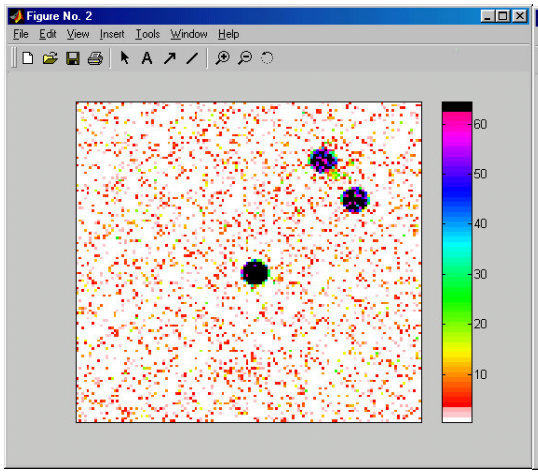
A simulation for the wide-field epi-illumination case was executed with the value of *zfocus* set to 1.0 μm, the images from which are shown in Figure 13. Because of defocus effects, larger structures than those in Figure 4 are seen. All molecules within the volume now receive the same excitation level and emit similar numbers of photons. However, only molecules at the center of the flowcell can be expected to generate well-focused images. Molecules at different depths result in unfocused images that blend into adjacent images. As time progresses, later frames show a reduction in intensity and structure due to photo-bleaching.



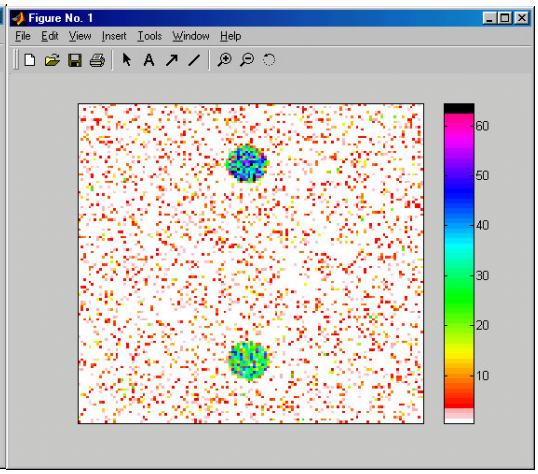
zfocus = 0.0



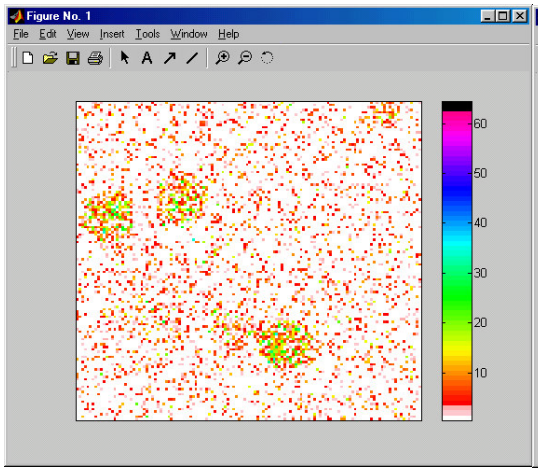
zfocus = 0.1



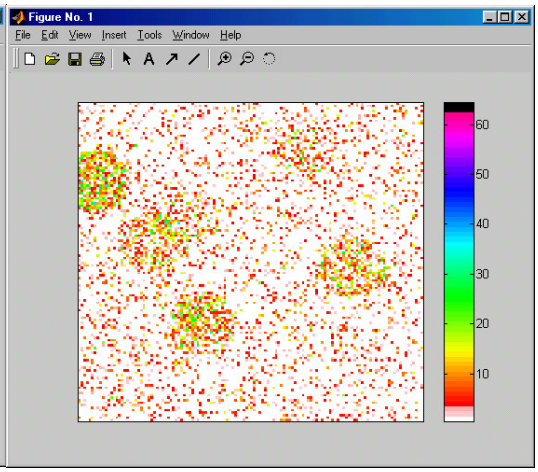
zfocus = 0.2



zfocus = 0.3

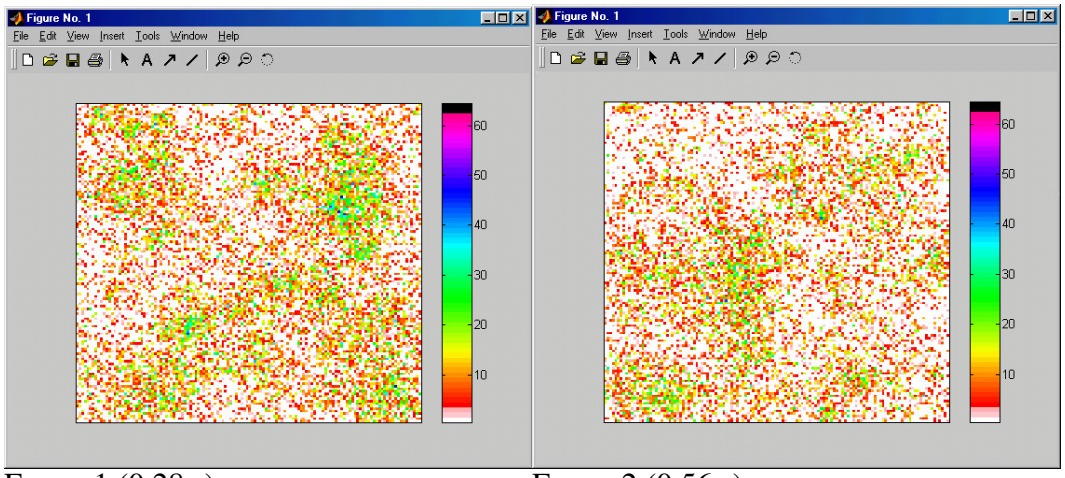


zfocus = 0.4

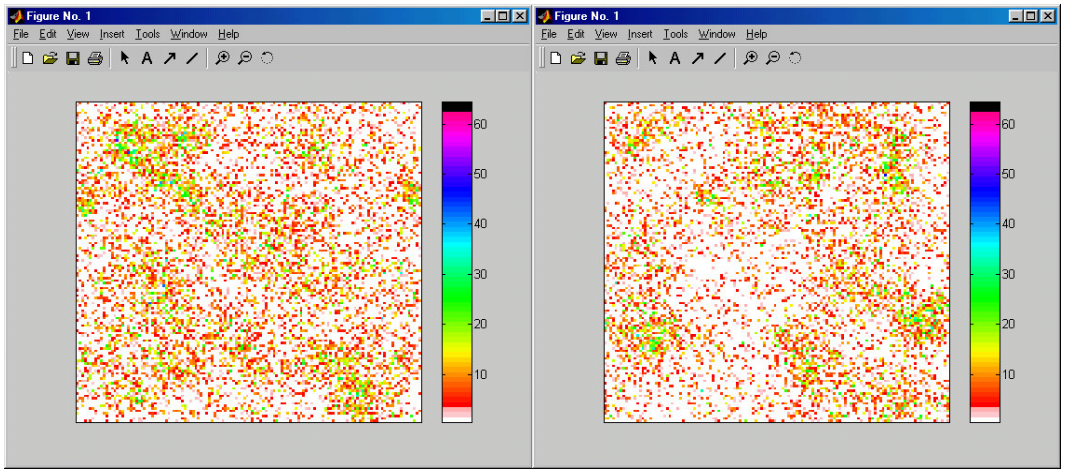


zfocus = 0.5

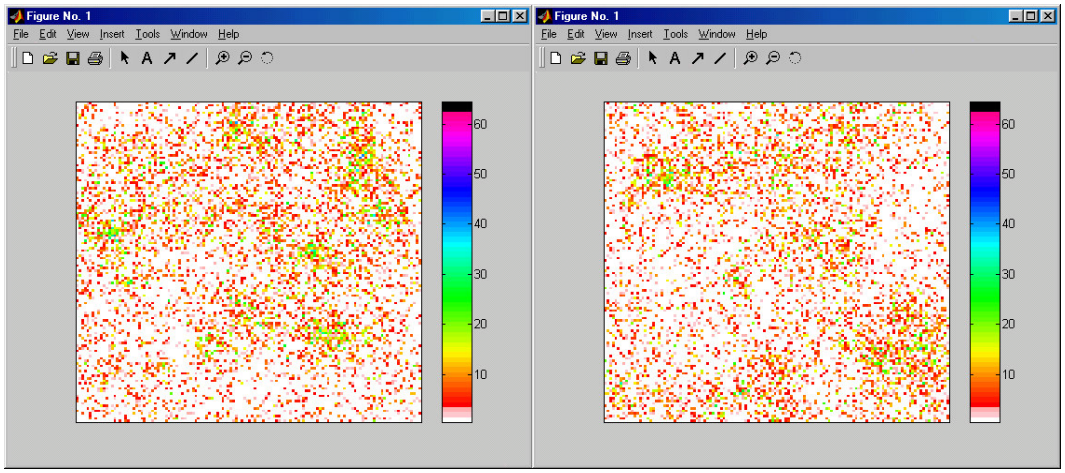
Figure 12: The effects of moving the plane of focus away from the substrate surface on which molecules stick.



Frame 1 (0.28 s) Frame 2 (0.56 s)



Frame 3 (0.84 s) Frame 4 (1.12 s)



Frame 5 (1.40 s) Frame 6 (1.68 s)

Figure 13: The effect of setting the evanescent wave penetration depth to infinity so that all molecules are equally illuminated.

The simulation results of Figures 12 and 13 indicate that it will be necessary to decrease the depth of the flowcell to reduce defocus effects in order to image all the free molecules within a flowcell.

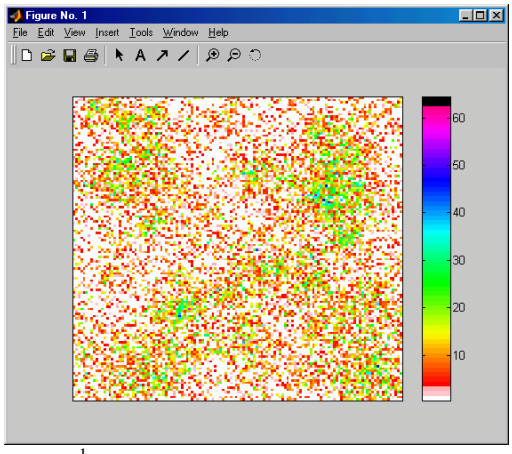
2.2.6. Addition of Motion to the Flowcell

To this point, the solution within the flowcell has been assumed stationary with no net flow, i.e., the molecules move only due to diffusion. In this section, constant translational motion in the y-direction with an adjustable flow speed {1} is added for each molecule {5}.

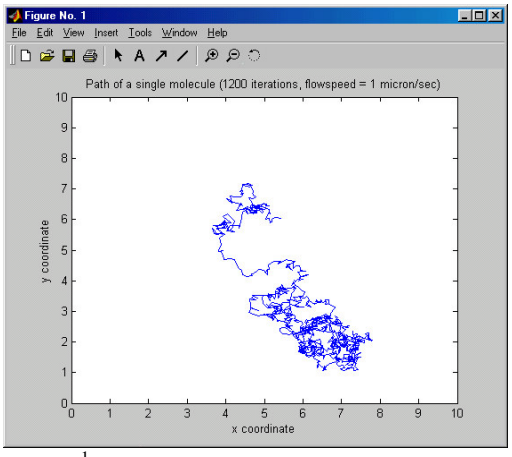
To correctly implement the flow, the spatial boundary conditions of the probe volume must also be altered. In all previous simulations, these boundary conditions are such that molecules simply bounce from all interfaces. However, with flow, molecules would be pushed toward one boundary and would congregate there as they constantly bounce off that boundary. Instead of molecules bouncing, periodic boundary conditions are used, which dictate that once a molecule leaves a boundary, a replica enters at the opposite boundary {8}.

Figure 14 shows a set of images from simulations with the flow and boundary condition code implemented, over a range of flow speeds. The parameters for this set of simulations are as in Section 2.1, but with the following conditions:

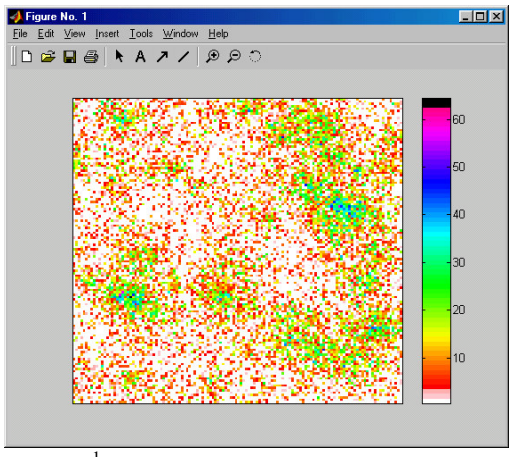
- No sticking
- Flow-cell depth = 2.0 μm
- Wide-field epi-illumination (with 3 ms pulsed excitation)



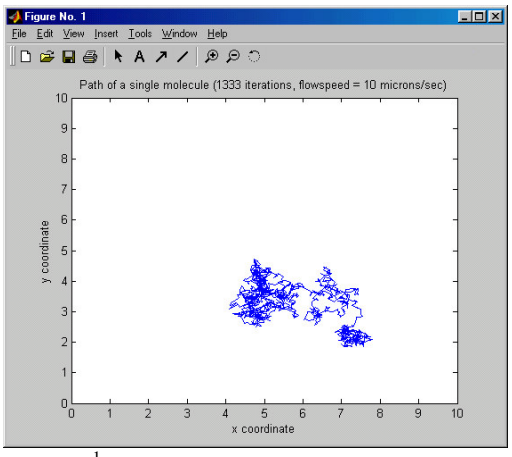
1 µm s^{-1}



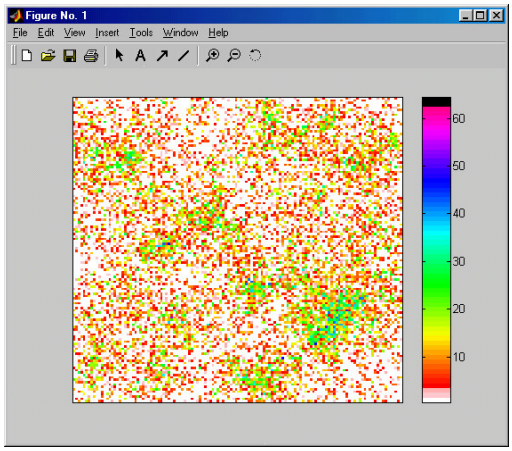
1 µm s^{-1}



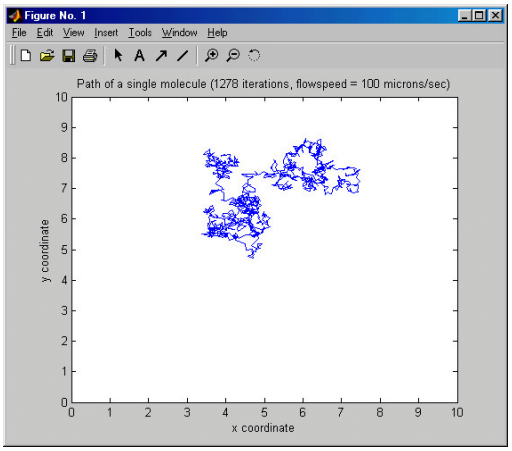
10 µm s^{-1}



10 µm s^{-1}

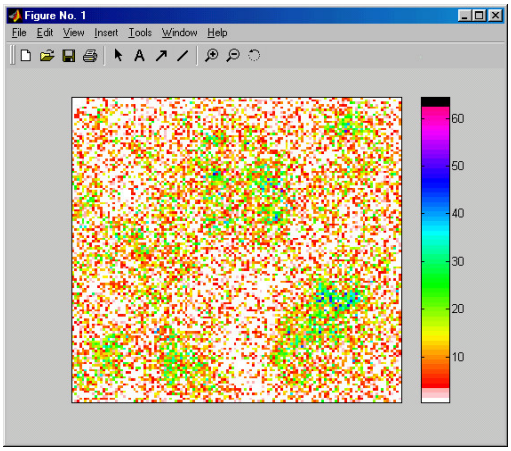


10^2 µm s^{-1}

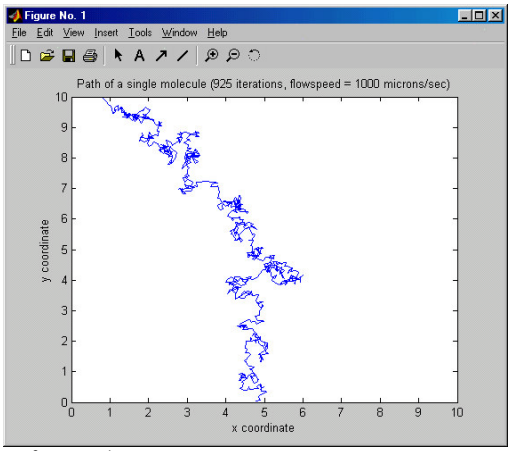


10^2 µm s^{-1}

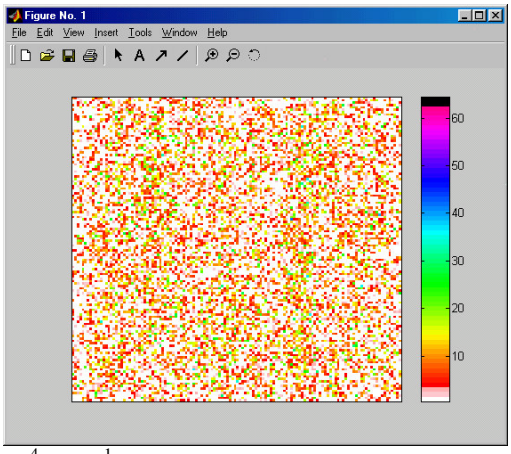
Figure 14: First frame images of the simulation for varying flow and the *x*-*y* paths of molecules: for flow rates of 1, 10 and 10^2 µm s^{-1}.



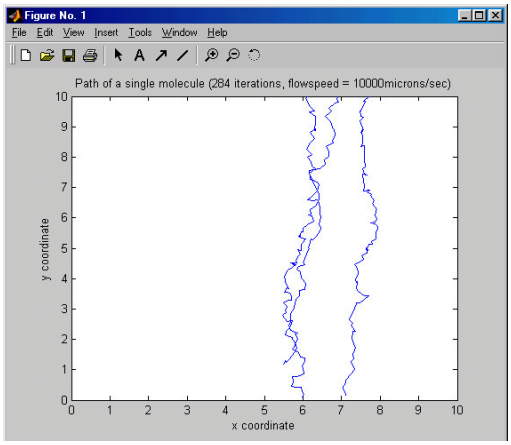
10^3 µm s^{-1}



10^3 µm s^{-1}



10^4 µm s^{-1}



10^4 µm s^{-1}

Figure 14: Continued: for flow rates of 10^3 and 10^4 µm s^{-1}.

- Objective focal plane, *zfocus* = 1 μm
- Flow rate = 1 - 10^4 μm s^{-1} (positive *y*-direction)
- Periodic boundary conditions on *x* and *y* positions

Each image of the set shows the first image frame of the simulation with a corresponding plot of the trajectory of a single molecule. The trajectory is obtained directly from *x* and *y* coordinates of the molecules in the simulation and not from analysis of the image.

From the plots of the trajectories, it is apparent that for distance scales of the order of 128 pixels (18.3 μm in object space) diffusion dominates for flow velocities < 10^2 μm s^{-1}, and flow dominates for flow velocities > 10^4 μm s^{-1}. The two contributions are approximately balanced for a flow speed of 10^3 μm s^{-1}, with molecules still discernable within the image.

The motion of molecules due to translation or diffusion compounds the defocus problems discussed in Section 2.2.5 and makes the imaging of moving molecules within a flowcell very difficult if not impossible.

2.3 Discussion

Monte Carlo simulations have successfully reproduced what is observed in experimental images and were used to help predict the outcomes of various experimental scenarios.

The first predictions show that if the laser exposure time is increased above 3 ms, the images begin to smear, and reach a near saturated image for exposure times of 280 ms. This is analogous to a camera shutter left open for a long time while taking a picture of a moving object.

Continuous laser exposure over a range of laser powers also results in smearing, but with different levels of intensity in the images. This would be detrimental for the imaging of freely diffusing molecules, as no discernable structure would be seen.

The sticking of molecules to the substrate surface was also successfully modeled. With reduced laser powers, it is found that molecules can be observed, although faintly, down to powers of 100 μW. Low irradiance may help to avoid effects such as photo-damage in live cell experiments if one is interested in the observation of only stuck-molecules. Should one wish to observe more intense images of stuck molecules, it should be noted that they can be seen just as well for continuous powers of 1 mW as for pulsed 325 mW irradiance, as used in the initial investigations.

Defocus effects were also added to the simulation. The increase in the spot size of the molecule image with distance from the focal plane, and the point at which defocus effects become dominant over diffraction were determined using Zemax optical design software.

Epi-illumination within a 2 μm deep flowcell was implemented to evaluate proposed DNA sequencing experiments. The focal plane was positioned at different depths to investigate defocus effects. Stuck molecules were imaged for increased focal depths with a correspondingly increased spot size, as one would expect. It was found that the intensity of the images at a distance of 1 micron away from the focal plane was reduced so much that direct observation of the molecule is impossible. To more clearly image molecules within the flowcell, the focal plane was set at the center of the flowcell. However, images of free molecules look smeared because of defocus effects.

Constant one-dimensional translational motion of the solution in the flowcell was modeled over a range of speeds. Translation effects on the molecule trajectories are

apparent for flow speeds above 10^3 μm s^{-1}, and some structure is still visible within the images. When the flow speed reaches 10^4 μm s^{-1}, the images lose all structure as the molecule moves too fast to create a sharp image for the laser exposure time of 3 ms. Shorter exposure times may help to pick out the molecule as it travels, but image quality may also suffer due to decreased intensity unless higher laser power were used.

The development of image processing procedures, which had been undertaken by a collaborator in parallel with the simulation modeling reported in this chapter, could possibly reassemble images from defocused or moving molecules. However the procedures were not available during the course of this research and such image processing is suggested for future work.

3. Analysis of Fluorescence Correlation Spectroscopy Data

3.1 Functional Form of the Autocorrelation Function (ACF)

The normalized autocorrelation function, $G(\tau)$, is defined as

$$G(\tau) = \frac{\int_{-\infty}^{+\infty} F(t)F(t+\tau)dt}{\left[\int_{-\infty}^{+\infty} F(t)dt\right]^2} = \frac{\langle F(t)F(t+\tau)\rangle}{\langle F(t)\rangle^2}, \tag{6}$$

where $F(t)$ is the fluorescence signal detected at time t. $G(\tau)$ correlates fluorescence between two moments in time, t and $t+\tau$, separated by a given time delay τ. For large time delays ($t \rightarrow \infty$), the fluorescence is uncorrelated and $G(\tau) \rightarrow 1$. References [24] and [25] show that

$$G(\tau) = 1 + \frac{\gamma}{N} G_D(\tau), \tag{7}$$

where γ is a proportionality constant dependent on the sample volume geometry, N is the mean number of molecules in the sample volume, and $G_D(t)$ is the contribution due to correlated fluorescence emission from a single molecule diffusing through the sample volume. Reference [25] gives an equation for $G_D(\tau)$

$$G_D(\tau) = \iint d^3\mathbf{r}_1 d^3\mathbf{r}_2 O(\mathbf{r}_1) C(\mathbf{r}_1 - \mathbf{r}_2, \tau) O(\mathbf{r}_2), \tag{8}$$

the one-dimensional form of which is

$$G_D^{(x)}(\tau) = \iint dx_1 dx_2 O(x_1) C(x_1 - x_2, \tau) O(x_2). \tag{9}$$

Here, $O(x_1)$ is proportional to the rate of fluorescence signal from a single molecule at x_1, $C(x_1 - x_2, \tau)$ is the probability density that the molecule initially at x_1 is found at x_2 after a time τ, and $O(x_2)$ is proportional to the rate of fluorescence from the molecule at x_2. In the expression for $G_D^{(x)}(\tau)$, the ensemble average of the correlated fluorescence emission is found by integrating over all possible values of x_1 and x_2.

For a three-dimensional volume element that is illuminated by a focused laser beam, and from which fluorescence is collected, the normalized volume profile is defined by

$$O(\mathbf{r}) = S(\mathbf{r})\Omega(\mathbf{r}). \tag{10}$$

whereby

$$\iiint O(\mathbf{r}) d^3\mathbf{r} = 1, \tag{11}$$

and $S(\boldsymbol{r})$ and $\Omega(\boldsymbol{r})$ are the illumination and fluorescence collection efficiency profiles respectively. $O(\boldsymbol{r})$ is approximated to be a three-dimensional ellipsoidal Gaussian, although $S(\boldsymbol{r})$ and $\Omega(\boldsymbol{r})$ depend upon the optics and are strictly not Gaussian, particularly along the optical axis, z. Recently, numerical calculation of $S(\boldsymbol{r})$ and $\Omega(\boldsymbol{r})$ and numerical evaluation of the integrals in equation (8) have been presented [25]. The three-dimensional form of $G_D(\tau)$ given in equation (8) is simply the product of three one-dimensional terms, as given in equation (9). Also, a three-dimensional Gaussian is simply the product of three one-dimensional Gaussians. For the x-component of $O(\boldsymbol{r})$, the normalized Gaussian is

$$O(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{x^2}{2\sigma^2} \right], \tag{12}$$

where the $1/e^2$ beam radius is $\omega_0 = 2\sigma$.

The function $C(\boldsymbol{r}_1 - \boldsymbol{r}_2, \tau)$ in equation (8) is found from the diffusion equation, i.e., Fick's Second Law. Most FCS papers (e.g. [26]) begin by application of Fick's Law to the concentration fluctuations, i.e., the difference between the concentration at a spatial location $\boldsymbol{r}$ and time t and the bulk-averaged concentration

$$\delta C(\mathbf{r},t) = C(\mathbf{r},t) - \langle C(\mathbf{r},t) \rangle , \tag{13}$$

to obtain

$$\frac{\partial}{\partial t}\delta C(\mathbf{r},t) = D\nabla^2 \delta C(\mathbf{r},t) . \tag{14}$$

However, the approach adopted here, which is appropriate for concentrations relevant to SMD, is to consider the probability density to find a single molecule at a particular point in space and time, $C(\boldsymbol{r},t)$. Application of Fick's Law now yields

$$\frac{\partial}{\partial t}C(\mathbf{r},t) = D\nabla^2 C(\mathbf{r},t) , \tag{15}$$

which describes the relationship between the spatial and temporal behavior of the probability density. As the spatial components of equation (15) are separable, for simplicity, only the one-dimensional case need be considered:

$$\frac{\partial}{\partial t}C(x,t) = D\frac{\partial^2 C(x,t)}{\partial x^2} . \tag{16}$$

For the initial condition that a molecule starts at the origin at $t = 0$, modeled by the Dirac Delta Distribution $C(x,0) = \delta(x)$, the solution of equation (16) takes the form of a Gaussian:

$$C(x,t) = \frac{1}{(4\pi Dt)^{1/2}} \exp\left[\frac{-x^2}{4Dt}\right] . \tag{17}$$

Substitution of (12) and (16) into (9) yields

$$G_D^{(x)}(\tau) = \frac{1}{2\pi\sigma^2}\frac{1}{(4\pi D\tau)^{1/2}}\int\int dx_1 dx_2 \exp\left\{-\left[\frac{({x_1}^2 + {x_2}^2)}{2\sigma^2} + \frac{(x_1 - x_2)^2}{4D\tau}\right]\right\}. \tag{18}$$

To perform the integration over x_1 first, complete the square in x_1 for the factor in the square parenthesis:

$$\left[\frac{({x_1}^2 + {x_2}^2)}{2\sigma^2} + \frac{(x_1 - x_2)^2}{4D\tau}\right] = \left(\alpha x_1 - \frac{\beta}{\alpha}x_2\right)^2 + \left(\alpha^2 - \frac{\beta^2}{\alpha^2}\right){x_2}^2, \tag{19}$$

where

$$\alpha^2 = \frac{1}{2\sigma^2} + \frac{1}{4D\tau} \tag{20}$$

and

$$\beta = \frac{1}{4D\tau}. \tag{21}$$

Hence

$$G_D^{(x)}(\tau) = \frac{1}{2\pi\sigma^2}\sqrt{\frac{\beta}{\pi}}\int dx_2 \exp\left[-\left(\alpha^2 - \frac{\beta^2}{\alpha^2}\right){x_2}^2\right]\int dx_1 \exp\left[-\left(\alpha x_1 - \frac{\beta}{\alpha}x_2\right)^2\right]. \tag{22}$$

The standard formula

$$\int_{-\infty}^{+\infty}\exp\left[-(ax+b)^2\right]dx = \frac{\sqrt{\pi}}{a}, \tag{23}$$

yields

$$G_D^{(x)}(\tau) = \frac{1}{2\pi\sigma^2}\sqrt{\frac{\beta}{\pi}}\int dx_2 \exp\left[-\left(\alpha^2 - \frac{\beta^2}{\alpha^2}\right){x_2}^2\right]\sqrt{\frac{\pi}{\alpha^2}}, \tag{24}$$

$$= \frac{1}{2\pi\sigma^2}\sqrt{\frac{\beta}{\pi}}\frac{\pi}{\sqrt{\alpha^2 - \beta^2}},$$

$$= \frac{1}{2\sqrt{\pi}} \frac{1}{\sqrt{\sigma^2 + D\tau}},$$

$$G_D^{(x)} = \frac{2}{\sqrt{\pi}\omega_0} \frac{1}{\sqrt{1+\tau/\tau_D}}, \tag{25}$$

where

$$\tau_D = \frac{\omega_0^{\,2}}{4D}. \tag{26}$$

For diffusion in three dimensions with a 3-D ellipsoid sample volume,

$$O(x, y, z) = \frac{1}{\pi\omega_0^{\,2}} \exp\left[-\frac{2(x^2 + y^2)}{\omega_0^{\,2}}\right] \frac{1}{\sqrt{\pi} z_0^{\,2}} \exp\left[-\frac{2z^2}{z_0^{\,2}}\right], \tag{27}$$

where ω_0 is the sample volume $1/e^2$ radius (i.e., the laser beam waist) and z_0 is the volume axial radius. The three-dimensional form of $G_D(\tau)$ in equation (8) is the product of three terms like those of equation (25) yielding

$$G_D(\tau) = \frac{8}{\pi^{3/2}\omega_0^{\,2} z_0} \left[1 + \frac{\tau}{\tau_D}\right]^{-1} \left[1 + \left(\frac{\omega_0}{z_0}\right)^2 \frac{\tau}{\tau_D}\right]^{-1/2}. \tag{28}$$

For a three-dimensional ellipsoidal Gaussian sample volume, the term γ in equation (7) may be taken as

$$\gamma = \frac{\pi^{3/2}\omega_0^2 z_0}{8}, \tag{29}$$

in which case N is the mean number of molecules in a volume $\pi^3\omega_0^{\,2}z_0$ [24].

Substitution of equations (28) and (29) in equation (7) yields

$$G_D(\tau) = 1 + \frac{1}{N}\left[1 + \frac{\tau}{\tau_D}\right]^{-1} \left[1 + \left(\frac{\omega_0}{z_0}\right)^2 \frac{\tau}{\tau_D}\right]^{-1/2}. \tag{30}$$

Equation (30) forms the basis for most of the investigations of this chapter. Note however that if two or more solution components with substantially different diffusion coefficients are present, equation (30) may be simply extended by addition of terms similar to the last one [24]. Section 3.6 presents an extension of equation (30) to account for transport of molecules by flow in addition to diffusion.

3.2 Dependence of the Shape of the ACF on the Parameters

Equation (30) for the normalized ACF can be rewritten as

$$G(\tau) = 1 + a_1(1 + a_2\tau)^{-1}(1 + a_2 a_3\tau)^{-1/2}, \tag{31}$$

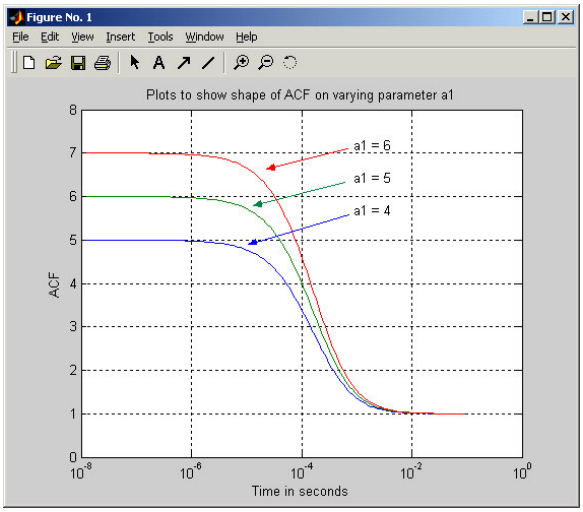
where

$$a_1 = \frac{1}{N} \quad \text{(a)}, \qquad a_2 = \frac{4D}{\omega_0^{\,2}} = \frac{1}{\tau_D} \quad \text{(b)}, \qquad a_3 = \frac{\omega_0^{\,2}}{z_0^{\,2}} = \frac{1}{\omega^2} \quad \text{(c)}. \tag{32}$$

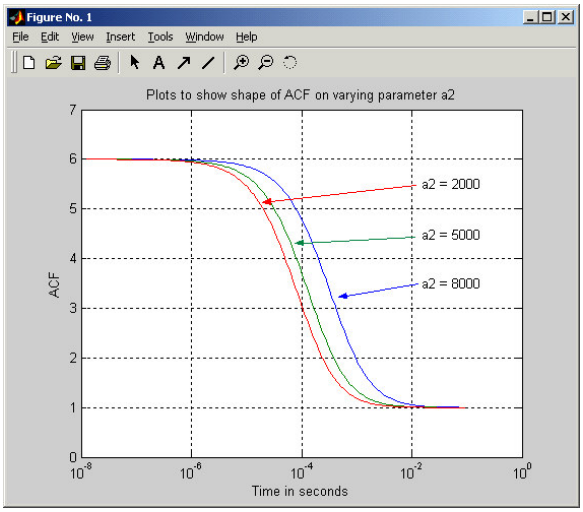
The role that the parameters, a_1, a_2 and a_3, play in the shape of the ACF can be found by plotting a series of curves where one parameter is varied and the other two fixed. Note that a logarithmic scale is generally used for τ to enable illustration of the dependence of $G(\tau)$ over a wide range of timescales.

Figure 15a shows a set of curves where only a_1, which determines the peak amplitude of the ACF, is varied. From equation (32a) it can be concluded that the higher the peak amplitude, the lower the mean number of fluorophores in the probe volume.

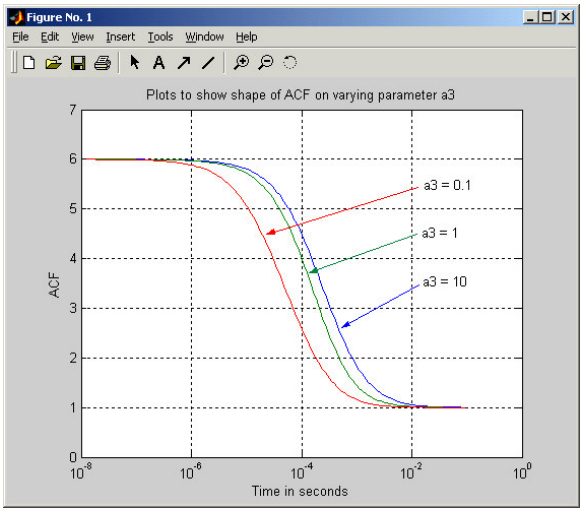
Variation of a_2 adjusts the temporal width of the ACF (Figure 15b). From equation (32b), the extent of this width dependent on two factors, ω_0 and D, and hence it is not possible



(a) Vary a_1 (a_2=4000, a_3=1)



(a) Vary a_2 (a_1=5, a_3=1)



(c) Vary a_3 (a_1=5, a_2=4000)

Figure 15: Dependence of the shape of the ACF on the a_1, a_2 and a_3 parameters.

to make an independent determination of both. Usually ω_0 is determined in a calibration experiment that uses a molecule with a well-known diffusion constant.

Similarly, the value of a_3 (Figure 15c), which depends on the geometry of the experimental set up (i.e., the factors ω_0 and z_0) can be determined in a calibration experiment and then held fixed. These parameters may now be adjusted within the ACF to fit any experimental data. Once that fit is achieved, actual experimental values for N, ω_0^2 and z_0^2, can be deduced, if D is known.

3.3 Curve-Fitting of Experimental ACF Data

Three sets of experimental data, collected by LI-COR using an FCS instrument constructed at UTSI as part of a research contract for LI-COR, corresponding to three solutions that contained different concentrations of a diffusing component, in this case TAMRA-SE, were investigated using the ACF fitting function. The fitting procedure was similar to that carried out in reference [27]. In that work the authors used Rhodamine 6G as the diffusing component and performed calibration measurements, as explained in Section 3.2, which yielded a diffusion time of 0.04 ms and hence a volume element with dimensions $\omega_0 = 0.2$ μm and $z_0 = 1.0$ μm.

Fits of the experimental data were performed with MATLAB and it's "fminsearch.m" function, which uses the Nelder-Mead simplex algorithm to minimize deviations between data and fit. Initial estimates for the parameters were obtained as follows:

a_1 takes the value of the first point in the data set minus 1, i.e., $a_1 = G(1) - 1$, where $G(i)$

refers to the *i*th value of the experimental ACF data set, which corresponds to a time τ in seconds.

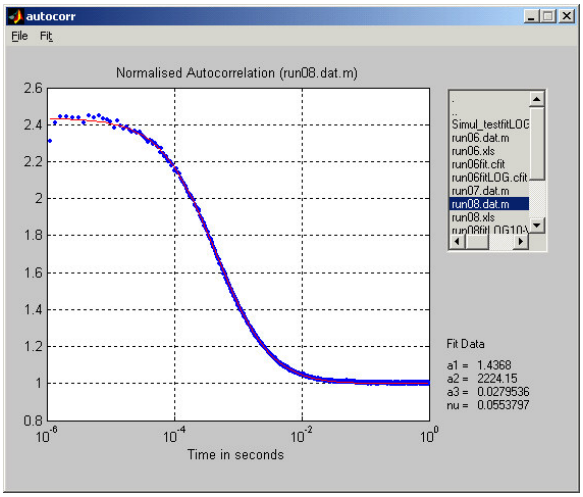
For a_2, a search is performed through the data set to find the first point, which has a value that falls below of one half of a_1. The time delay of this point, $\tau_{1/2}$, is then used to obtain $a_2 = (4^{1/3} - 1)/\tau_{1/2}$ as it's initial value.

a_3 uses the calibrated values of ω_0 and z_0 so that $a_3 = 0.04$ μm initially. For a reasonable fit, this value is expected to vary by a small amount, if at all.

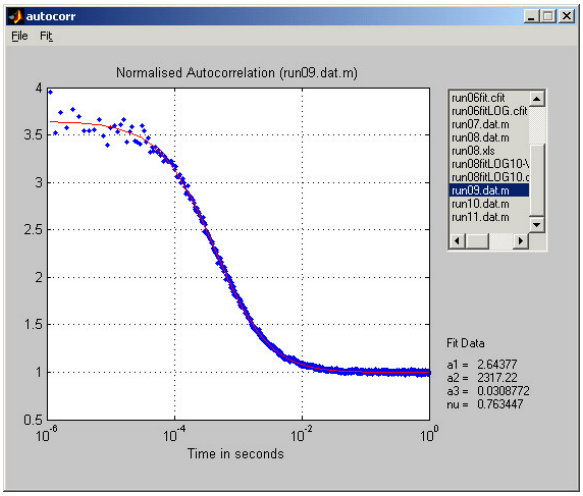
Execution of the "fminsearch" function, with these initial values and the experimental data set as input, results in the output of the values of a_1, a_2 and a_3 that give the best fit to the data, and the parameter v, which is the square of the unweighted deviations between data and fit as used in the fitting process. The fits to the three sets of experimental data are shown in Figures 16a, b & c respectively and the best-fit parameters are shown in Table 2.

The a_1 values do indeed vary for different concentrations as expected. However, simply using equation (32a) would not give the observed value. This is because there are other parameters wrapped up with a_1 such as the size of the probe volume [24] and the amount of measurable background photons present in the experiment [12] as well as numerical constants such as Avogadro's number ($N_A = 6.022\times10^{23}$ mol^{-1}).

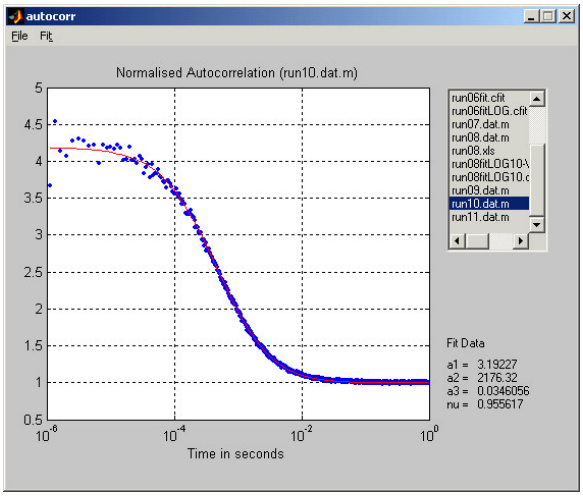
The a_2 values are similar, with the average being $a_2 = D/4\omega_0^2 = 2240$ s^{-1}. As the diffusion constant of a molecule, D, is approximately proportional to the cube root of the molecular weight, then D for TAMRA-SE should be approximately equal to that of Rhodamine 6G,



(a) Concentration = 1.25×10^{-11} M



(b) Concentration = 5.0×10^{-11} M



(c) Concentration = 1.25×10^{-12} M

Figure 16: Experimental ACF data for TAMRA-SE with various sample concentrations. The data (blue) is fit (red) the using ACF functional form (equations 27 & 28) with fit values for a_1, a_2 and a_3.

Table 2: Values of the three parameters that give the best fits to data using different molecular concentrations of TAMRA-SE.

	a_1	a_2	a_3
TAMRA-SE 1.25×10^{-11} M	1.4368	2224.2 s-1	0.028
TAMRA-SE 5×10^{-12} M	2.6438	2317.2 s-1	0.0309
TAMRA-SE 1.25×10^{-12} M	3.1923	2176.3 s-1	0.0346

$D = 2.8\times10^{-6}$ cm^2 s^{-1}. Solving equation (32b) for ω_0 gives a value of 0.177 μm for the beam waist, similar to that quoted by [27].

The value of a_3, although different for each data set, still remains near the initial value of 0.04 μm.

It can be concluded that the derived functional form and the fitting method achieve expected results for the extraction of kinetic information from experimental data, which promotes confidence should this method be used in other experiments.

3.4 MATLAB GUI for ACF Curve-Fitting

For the previous investigations, most of the commands required to read data, process the fit and output the parameters were initially performed via the command line of MATLAB. This proved to be not only tedious but also time consuming and open to user error. Thus a MATLAB program with graphical user interface (GUI), named 'autocorrp' (Appendix 2), was written to perform all aspects of the fitting process, including forming a plot of the data and the best fit, and the values of the fit parameters.

Figures 16 a, b & c from the previous section show plots produced by the GUI, which contains imported ACF data (blue dots), the best fit (red curve), and the respective parameters (lower right). The process to fit experimental data can be performed by carrying out the following steps:

1) Conversion of the ACF into Matlab format and import into the GUI environment:

   An experiment yields the normalized ACF as pairs of values of (τ, $G(\tau)$) listed in a text file that has a name with the extension '.dat' (e.g. filename.dat). The listbox

to the right of the GUI shows the files in the current directory. Find the directory that contains the raw data to be converted. Double-click on the file to create a new file with the original name, but with '.m' appended (e.g. filename.dat.m). The data contained in this file is then plotted on a graph displayed at the center of the GUI.

Alternatively, click on 'File' in the toolbar, then 'create' and choose the '.dat' file to convert. The file will be converted and the data plotted.

2) To plot an already existing '.m' data file:

   Data that has already been converted can be imported into the GUI environment. To achieve this, find the directory in which the '.m' file is contained then double-click on the required filename. The data is then plotted to the graph.

   Another way is via the toolbar. Click 'File' then 'Plot'. The user can select a file, which is then plotted.

3) To create a fit to the data:

   Once a data set has been imported and plotted, a fit may be performed.

   Click on 'Fit' in the toolbar, then 'Gfit1'. This performs a single-component ACF fit to the data, i.e., a fit of the ACF to equation (31). The curve of the fit is shown in red with the values of the fit parameters appearing towards the bottom-right of the GUI screen.

4) Other functions:

   The 'File' menu in the toolbar also contains the functions:

   Print – Prints the figure including graph and data

Close – Removes the current graph and respective data allowing new data to be analyzed

Exit – Exits the GUI

The GUI was tested and found to perform well allowing the user rapid display and fitting of experimental data within a user-friendly environment.

3.5 FCS Experiment to Study the Influence of Transcription Factor II on the Stable Binding of TATA-Binding Protein (TBP)

Current ideas in fundamental molecular biology [28] suggest that, although DNA seems to control the order of amino acids in proteins, it does not mean that DNA directly controls the synthesis of the proteins themselves. This follows from the understanding that DNA is confined to the nucleus of eukaryotic cells (cells with a nucleus) whereas protein production takes place within the cell cytoplasm, i.e., outside and surrounding the nucleus. Thus there is expected to be a bridging molecule to pass genetic information from the DNA to the proteins.

The molecule proposed for this function is RNA mainly because it has similar structure to DNA, and can easily be synthesized from a DNA template.

The progression of passing the DNA information to the proteins follows a suggested pathway of two processes given by:

DNA → RNA → Protein.

The first process, called transcription, forms the basis of the presently considered experiments, and the second is called translation.

Different classes of RNAs exist; the ones being relevant to the current discussion are called messenger RNAs (mRNAs), which are transcribed from the protein-coding DNA by a structure called an RNA polymerase II transcription complex. However, for this structure to successfully bind to the DNA strand, there must already be attached other specific proteins called transcription factors, and once RNA polymerase II is bound, further proteins must subsequently bind before the transcription process can begin.

From in vitro studies, the currently accepted chain of events from an empty DNA chain to one where the full protein complex is attached so that transcription can occur is illustrated in [28]. The DNA strand contains a sequence of nucleotides TATAA, called the TATA box, which occurs at around 25-30 nucleotides away from the position where transcription should begin. The TATA box is recognized by a protein complex called TFIID, which is made up of a TATA-binding protein (TBP) and TBP-associated factors (TAFs). However the binding may be short lived unless it is followed by binding of another molecule, TFIIB, which stabilizes the complex. The polymerase may then bind, along with a further protein TFIIF. The resultant complex is then completed with the binding of TFIIE and TFIIH, after which the transcription process may begin.

In vitro experiments were undertaken [29] to study the stabilization of TBP binding by the presence of TFIIB. The experiments begin with a solution containing DNA strands acceptable for transcription, along with TBP that was fluorescently labeled with Alexa-590, for which the diffusion time is known. Different concentrations of TFIIB are added to the solution to promote TBP binding to the DNA. Binding could be identified by

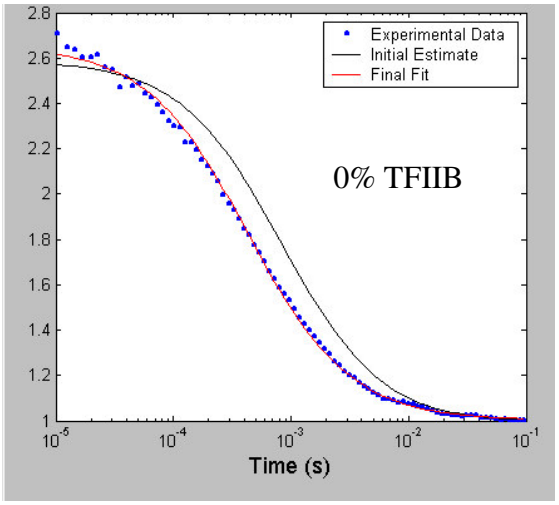
the use of FCS to observe a small change in the diffusion time ($\tau_D = \omega_0^2/4D$) of the labeled TBP, which becomes part of a larger complex with slower τ_D. As the diffusion coefficient, D, is approximately dependent on the cube root of the molecular weight of the complex, the difference in diffusion times between free and bound TBP is small but measurable.

Fluorescence from the experiments was collected for three concentrations of TFIIB, 0%, 0.07% and 0.2%, for collection times of about two minutes. Curve-fitting of the ACFs from the experiments was performed using the previously described GUI (Figures 17 a, b & c respectively). It can be seen that as the concentration of TFIIB was increased, the diffusion time of the labeled complex increases from τ_D = 430 μs for 0% TFIIB, to τ_D = 437 μs for 0.07% TFIIB, up to τ_D = 458 μs for 0.2% TFIIB. These results confirm that the TBP/TATA box bind is strengthened by the presence of TFIIB. Control experiments exhibited no change. These FCS experiments have thus directly supported the theories of the early stages of RNA polymerase II transcription complex formation.

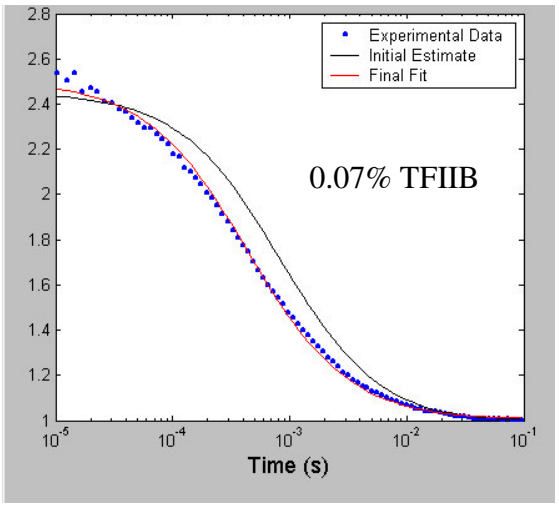
3.6 Extension of the Functional Form of the ACF to Include Flow

Flow may be used in FCS to decrease the effects of photo-bleaching and to increase the number of molecules that are probed in a short duration experiment.

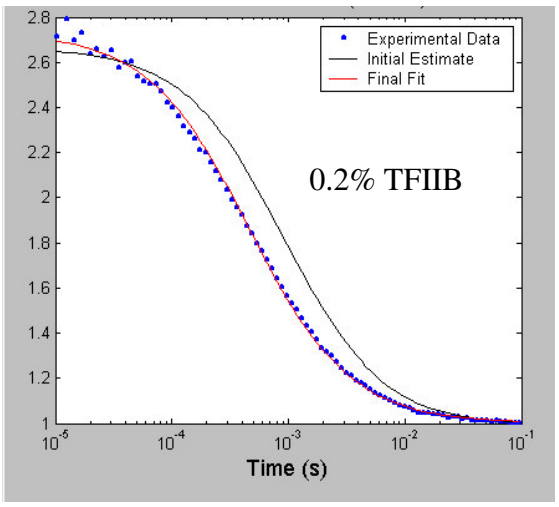
The derivation of the functional form of the ACF when flow is present begins with equation (9). However, as opposed to the diffusion-only equation (16), when a constant flow in the x-direction, with velocity v, is taken into consideration, the one-dimensional



(a) $\tau_D = 430\ \mu s$



(b) $\tau_D = 437\ \mu s$



(c) $\tau_D = 458\ \mu s$

Figure 17: The ACF from FCS experiments to study the influence of TFIIB on the stable binding of TBP to DNA. The plots correspond to different TFIIB concentrations.

diffusion equation becomes:

$$\frac{\partial}{\partial t}C(x,t) = D\frac{\partial^2 C(x,t)}{\partial x^2} - v\frac{\partial C(x,t)}{\partial x} \tag{33}$$

and the solution in equation (17) becomes

$$C(x_1 - x_2, t) = \frac{1}{(4\pi Dt)^{1/2}}\exp\left[\frac{-(x_1 - x_2 - vt)^2}{4Dt}\right]. \tag{34}$$

With the normalized volume profile as given in equation (12), and inserting equations (12) and (34) into equation (9), the integration can be performed, which results in the functional form for the ACF with flow (see Appendix 3):

$$G_D(\tau) = 1 + \frac{1}{N}\left[1 + \frac{\tau}{\tau_D}\right]^{-1}\left[1 + \left(\frac{\omega_0}{z_0}\right)^2\frac{\tau}{\tau_D}\right]^{-1/2}\exp\left[-\frac{1}{\left(1 + \frac{\tau}{\tau_D}\right)}\left(\frac{\tau}{\tau_F}\right)^2\right], \tag{35}$$

$$= 1 + a_1(1 + a_2\tau)^{-1}(1 + a_2a_3\tau)^{-1/2}\exp\left[-\left(\frac{\tau}{\tau_F}\right)^2\frac{1}{(1 + a_2\tau)}\right], \tag{36}$$

with a_1, a_2 and a_3 still defined by equations (32a, b & c) respectively. The flow contribution is contained within the exponential where the characteristic flow time is

$$\tau_F = \frac{\omega_0}{v_F}, \tag{37}$$

where v_F is the flow velocity.

3.7 Dependence of the Shape of the ACF on Flow Velocity

The form function now under consideration not only has a Lorentzian contribution as in equation (31), but also an exponential component where the exponents comprise of both

the diffusion and flow components. As the Lorentzian has been investigated in Section 3.2, the variables that contributed to that shape are left constant while the contribution due to flow is studied.

A set of plots was created with MATLAB to see how the shape of the form factor changed due to adding linear flow. The Lorentzian parameters, set to be $a_1 = 5$, $a_2 = 4000$ s^{-1} and $a_3 = 1$, were held fixed while the flow parameter was varied. As the value of a_2 is equivalent to a diffusion time of $\tau_D = 2.5\times10^{-4}$ s, for any effects of flow to be seen, the flow time, τ_F should be less than τ_D, i.e., the particles would now move faster through the volume element due to the effects of flow as opposed to only diffusion. Figure 18 shows the set of plots for flow times of 10^{-3} s, 10^{-4} s, 10^{-5} s, 10^{-6} s. It can be seen that for $\tau_F = 10^{-3}$ s, the curve (in red) is the same as that for if the movement was solely due to diffusion. Subsequent decrease of τ_F increases the role of the exponential term causing the 'drop off' to occur more to the left of the curve.

3.8 Experiments and Simulations that Include Flow

FCS experiments involving the use of flow are currently under investigation at UTSI. Many similar experiments are reported in the literature, including those described below, where different means were used to induce flow:

- Pressure driven flow using a syringe pump [26] and an elevated reservoir [30];
- Application of a voltage to negatively charged molecules to induce flow through a pipette [31];
- Application of a voltage to induce electro-osmotic flow [32].

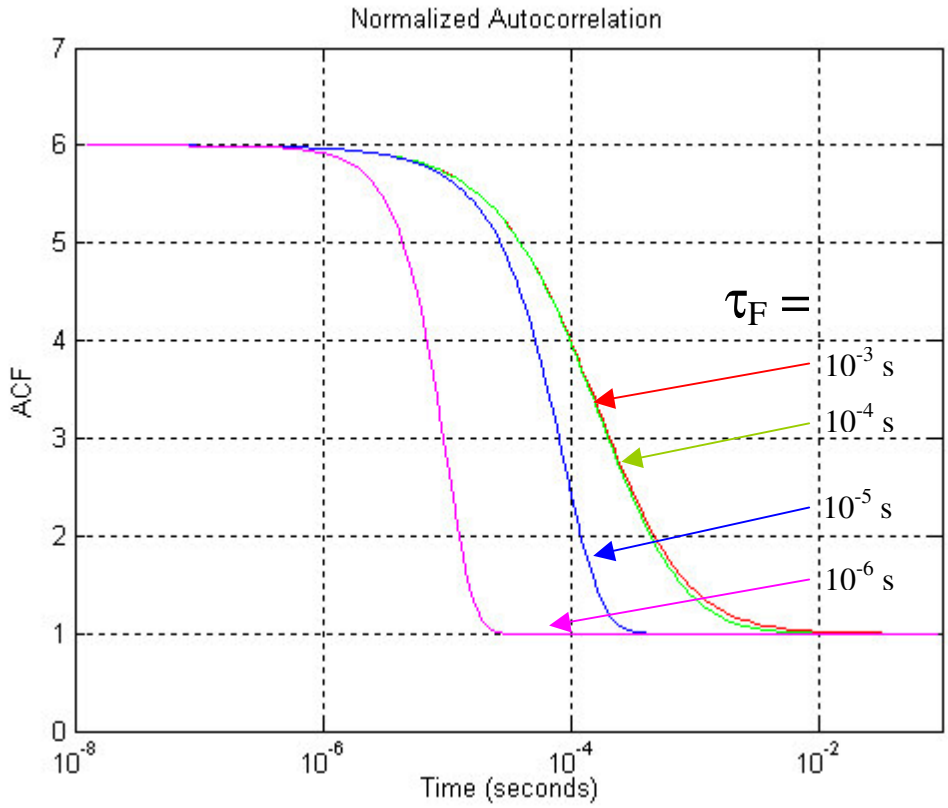


Figure 18: A set of plots to show the shape of the flow inclusive ACF form factor for varying flow rates.

The Zeiss Confocor 2 FCS instrument operated by collaborators at Abbott Laboratories enables the sample stage to be translated laterally with respect to the focused laser beam at speeds up to 24 mm s^{-1}. This motion is equivalent to molecules passing through the probe volume as if carried by a flow.

A Monte Carlo simulation of FCS has been developed at UTSI in LABVIEW and C [12], and has been applied to the Zeiss Confocor 2 set up. The parameters used in the simulation are listed in Table 3. A view of the corresponding sample volume produced by the simulation is shown in Figure 19. The ACFs generated by the simulation are used in this thesis research as input data for curve-fitting as described in the following section.

3.9 Curve-Fitting of the ACF with the Flow Model

The GUI described in Section 3.4 only fits zero-flow FCS data, whereas the fitting function to use with flow is given by equation (36). However, to run a simulation similar to that described above and in Table 3 but with no flow produces an ACF that can be fitted with the GUI in order to first obtain the a_1, a_2 and a_3 parameters. A fit to the flow simulation ACF can then be attempted using equations (36) with a_1, a_2 and a_3 held fixed while τ_F is varied.

Figure 20 shows the GUI with the zero-flow simulation data in blue and the best fit in red, with the three best-fit parameters $a_1 = 4.2$, $a_2 = 3117\ s^{-1}$, $a_3 = 0.437$. The plot of the ACF from the flow simulation in Figure 21 has a shape similar to the functions shown in Figure 18. However, it is clear that the amplitude is reduced compared to that shown in Figure 20, whereas the functional form derived in equation (36) predicts no such change.

Table 3: Parameters used in the simulation of the Confocor 2 experiment.

Parameter	**Value**	**Dimension**
Collection Time	10	s
Numerical Aperture (*NA*)	1.2	
Magnification	63	
Pinhole Radius	35	μm
Depth-of-Collection* (z_0)	0.27	μm
Laser Wavelength (λ)	488	nm
Laser Beam Waist** (ω_0)	0.26	μm
Laser Excitation Power	25	μW
Background Counts	125	s^{-1}
SPAD Quantum Efficiency	0.65	
SPAD Dead Time	35	ns
Molar Absorptivity	6.68×10^{4}	cm^{-1} $mole^{-1}$
Fluorescence Quantum Efficiency	0.9	
Bleaching Quantum Efficiency	10^{-5}	
Triplet Crossing Efficiency	10^{-4}	
Fluorescence Lifetime	3.7	ns
Triplet Lifetime	40	μs
Molar Concentration	6.5×10^{-10}	moles $liter^{-1}$
Diffusion Coefficient (*D*)	5.5×10^{-7}	cm^2 s^{-1}
Linear Velocity (v_F)	24	mm s^{-1}
Data Analysis Clock Tick	50	ns

* z_0 in equations (30) and (35)

** ω_0 in equations (30) and (35)

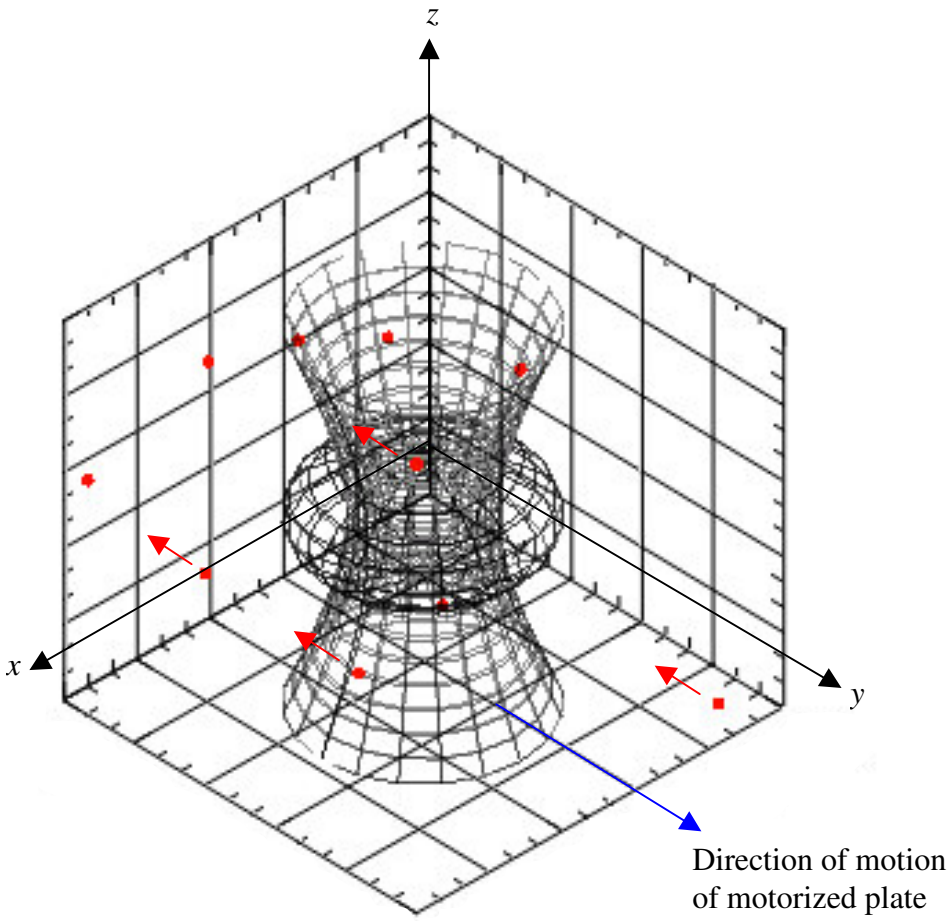


Figure 19: Schematic of the set up used by a Monte Carlo simulation to simulate the steady flow of molecules through a volume element. The volume element is defined by the laser beam width and the depth-of-collection.

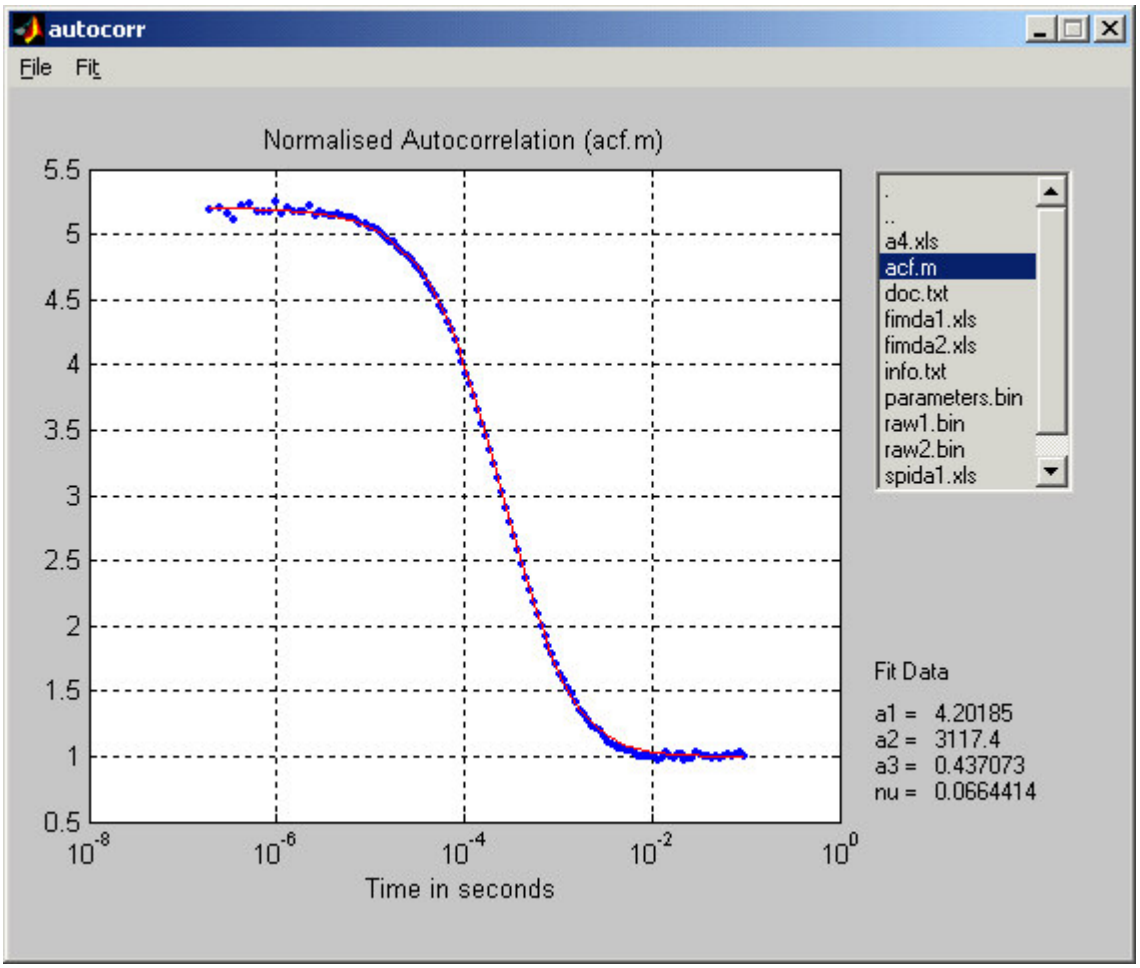


Figure 20: Fitted autocorrelation function for simulation data with no flow. From this fit, the a_1, a_2 and a_3 parameters are obtained.

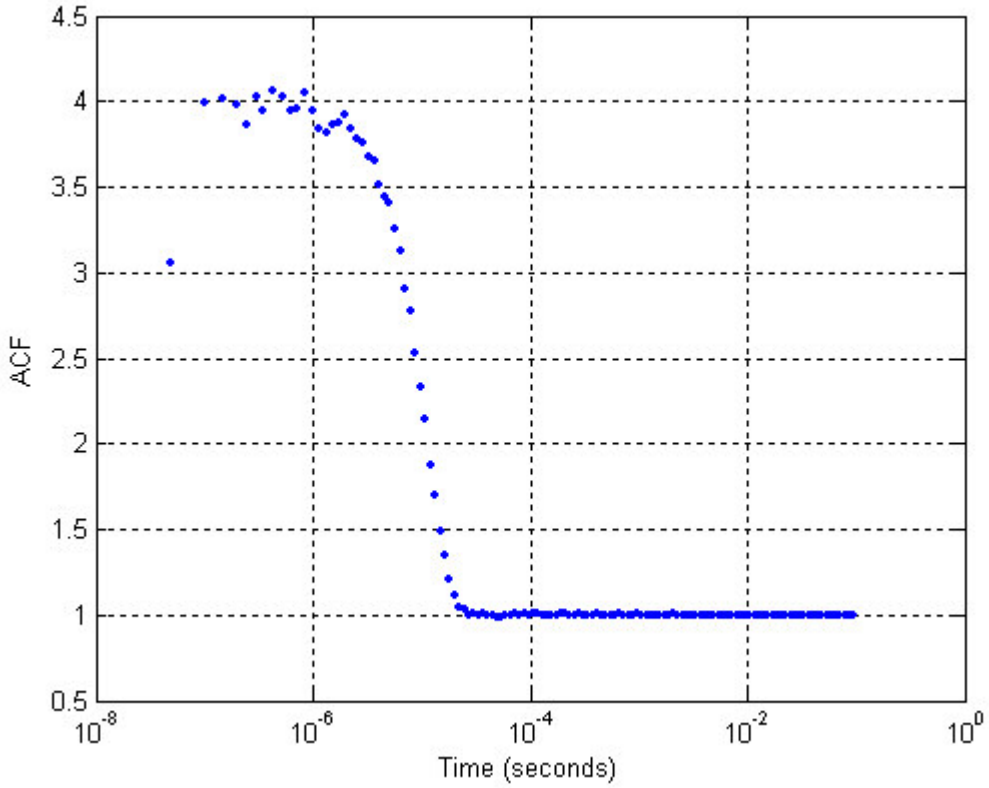


Figure 21: Autocorrelation function for anomalous simulation data.

Further investigation led to the discovery of a bug within the simulation, which effectively causes the sample concentration to vary with flow speed. On fixing the bug and re-running the simulation, the amplitude of the ACF was found to be independent of flow speed, consistent with the theoretical derivation of Section 3.6 and Appendix 3. The new ACF and the fit to equation (36) are shown in Figure 22 (blue and red curves respectively). The best-fit parameter for the characteristic time flow is $\tau_F = 1.1\times10^{-5}$ s, which agrees well with the actual value of 1.08×10^{-5} s determined from equation (37) and the simulation values listed in Table 3.

In conclusion, the ACF flow model results in a successful fit to the simulated experiment.

3.10 Discussion

Fluorescence Correlation Spectroscopy is a very useful tool for the determination of the kinetic properties of diffusing molecules. These properties can be extracted by fitting the theoretical autocorrelation function (ACF) to an experimentally obtained curve constructed by correlation of all photons with one another. The theoretical fitting function was derived and a program, which adopted a GUI format, was created to perform the curve-fitting process.

The program was used to successfully fit experimental data, which showed that TFIIB influences the binding of TBP to the TATAA box within a DNA strand during the process of creating a polymerase complex to transcribe DNA to RNA.

Processes such as high-throughput pharmaceutical drug screening would make use of FCS to characterize binding of library compounds to target proteins and would require as

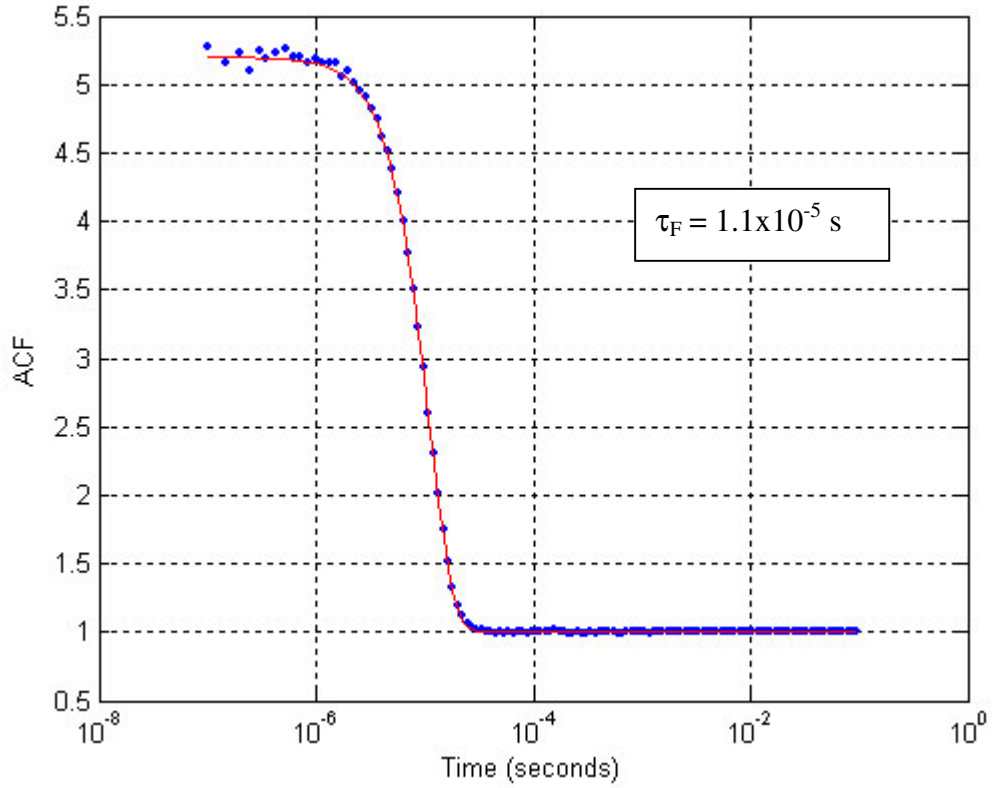


Figure 22: Fitted autocorrelation function for correct simulation data. The flow component is fitted by varying τ_F and keeping the a_1, a_2 and a_3 parameters fixed.

many molecules as possible to be studied in a short time. A method to do this is to introduce flow to the FCS system, to cause molecules to pass more quickly through a probe volume. Data from a simulation of FCS that included flow was used with a modified version of the fitting function that includes the presence of flow, and a successful fit was made.

Future work will be to extend the ACF analysis to consider different fluorescent species within a solution and to add this case and the case for flow into the fitting program.

4. Conclusions

4.1 Summary

Single-molecule detection techniques have been applied to study situations, which may be useful in many areas of biology, such as pharmaceutical drug screening, and which increase the potential for the development of rapid DNA sequencing. The work carried out as reported in this thesis studied aspects within the fields of single-molecule imaging and fluorescence correlation spectroscopy (FCS).

Monte Carlo simulations were used to model observations and predict results within single-molecule imaging (SMI) experiments to aid understanding of what is seen in experimental images and to plan future work. Increased laser exposure within the simulation, even with a decreased laser power, proved detrimental to the quality of images of freely diffusing molecules due to smearing effects. Simulated molecules stuck to the surface of the flowcell were found to be easier to image than those that diffused freely. Once defocus effects were introduced, image spot sizes were seen to be larger for those found outside the focal plane of the objective lens. When epi-illumination was adopted in place of TIR-illumination more molecules were imaged throughout the flowcell, but defocus effects caused molecules away from the focal plane to be formed into one larger image. Effects due to flow were only seen for flow velocities of 1000 μm s^{-1} and greater. Higher flow velocities resulted in image smearing.

A study of the fitting function of the autocorrelation function used in FCS methods was made. A MATLAB program using a GUI was written to perform a fit to experimental data using the fitting function. The GUI was used to fit autocorrelation data from a

protein binding experiment, which yielded useful data used to study processes with DNA to RNA transcription. The model was then extended to include molecule motion due to translation from a uniform flow as well as diffusion. This was used to fit data produced by a Monte Carlo simulation of a potential experiment with a flowing medium, which may be useful in high throughput screening of drug binding [33] or other rapid detection work.

4.2 Other Developments

Since beginning the research work contained within this thesis, advancements within the many fields of SMD continue to be made, some of which is relevant to the work described in the previous chapters.

Section 1.2 mentioned the application of SMI experiments, such as those simulated in Chapter 2, to DNA sequencing methods that prompted the research contained within this thesis. Although the Sanger method has been successfully applied to DNA sequencing in the Human Genome Project, it has limitations in cost, speed and sensitivity. Many faster and cheaper ways of whole-genome sequencing are now under investigation, many of which adopt single-molecule detection methods [34], with some experiments able to study individual DNA strands with single base resolution [35].

The improvement of sensitivity of signal data within FCS experiments has meant that the data integration time of such experiments has decreased from many hours required by early experiments [18] to the order of seconds. The length of the integration time has a large influence on the statistical accuracy of FCS, which has been an issue since it's

inception [36]. Statistics in FCS have since been extended to include other factors such as geometry of the laser beam [37]. Calculations of the variance in the ACF have suggested there exists a bias in the ACF data for large lag times [38], although computer simulations using symmetrical normalization methods of building the ACF have not yet provided evidence of this [12]. Further, the simulations have shown that since the points within the ACF are related, the variance of the points cannot be considered independently.

Data collection methods using the experimental apparatus at UTSI have been recently extended. Two counters for capturing single-photon event data are in use: National Instruments PCI-6602 (80 MHz) and Picoquant Time Harp 200 (20 MHz) [39]. Data is collected and analyzed using software custom built using C-programming language and the National Instruments Labview [40] graphical development environment.

References

List of References

[1] Hirschfeld T., "Optical Microscopic Observation of Small Single Molecules", *Appl. Opt.*, **15**, 2965-2966 (1976)

[2] Dovichi N.J., Martin J.C., Jett J.H., Trukula M., Keller R.A., "An approach to single-molecule detection by laser-induced fluorescence", in *Laser-Based Ultrasensitive Spectroscopy and Detection*, **Proceedings of SPIE Vol. 426**, 71-73 (1983)

[3] Shera E.B., "Single Molecule Tracking", *U.S. Patent*, **4793705** (1988)

[4] Li L.Q., "Single Molecule Detection in Solution", *PhD Thesis*, University of Tennessee, Knoxville (1994)

[5] Li L.Q., Davis L.M., "Rapid and Efficient Detection of Single Chromophore Molecules in Aqueous Solution", *Appl. Opt.*, **34**, 3208-3217 (1995)

[6] Bunfield D.H., "Simulation of a Single-Molecule Detection Experiment", *MS Thesis*, University of Tennessee, Knoxville (1997)

[7] Bunfield D.H., Davis L.M., "Monte Carlo Simulation of a Single-Molecule Detection Experiment", *Appl. Opt.*, **37**, 2315-2326 (1998)

[8] Service R.F., Moerner W.E., Orrit M., Weiss S., Gimzewski J.K., Joachim C., Mehta A.D., Rief M., Spudich J.A., Smith D.A., Simmons R.M., "Single Molecules", *Science*, **283**, 1667-1695 (1999)

[9] Ambrose W.P., Goodwin P.M., Nolan J.P., "Single Molecule Detection With Total Internal Reflection Excitation: Comparing Signal-To-Background and Total Signals In Different Geometries", *Cytometry*, **36**, 224-231 (1999)

[10] Sun Y., "Analysis of Ultrasensitive Fluorescence Experiments", *PhD Thesis*, University of Tennessee, Knoxville (2000)

[11] See webpage at http://www.mathworks.com/products/matlab/

[12] Davis L.M., Ball D.A., Williams P.E., Matayoshi E.D., Swift K.M., "Dealing with reduced data acquisition times in fluorescence correlation spectroscopy for HTS applications", in Microarrays and Combinatorial Technologies for Biomedical Applications, *SPIE Proc.* **4966**-20, ISBN 0-8194-4766-8 (2003)

[13] Ishikawa M., Hirano K., Hayakawa T., Hosoi S., Brenner S., "Single Molecule Detection By Laser Induced Fluorescence Techniques With A Position Sensitive Photon Counting Apparatus", *Jpn. J. Appl. Phys.* **33**, 1571-1576 (1994)

[14] Dickson R.M., Norris D.J., Tzeng Y.L., Moerner W.E., "Three Dimensional Imaging of Single Molecules Solvated in Pores of Poly(alrylamide) Gels", *Science*, **274**, 966-969 (1996)

[15] Williams J.G.K., "System and Apparatus for Nucleic Acid Sequencing of Single Molecules By Polymerase Synthesis", *U.S. Patent*, **US 20020115076A1**, (2002)

[16] Davis L.M., Parker W.C., Ball D.A., Williams J.G.K., Bashford G.R., Sheaff P., Eckles R., Lamb D.T., Middendorf L.R., "Imaging of Single-Chromophore Molecules in Aqueous Solution Near a Fused-Silica Interface", in *Multiphoton Microscopy in the Biomedical Sciences*, Ammasi Periasamy, Peter T.C. So, Editors, **Proceedings of SPIE Vol. 4262**, pp. 301-311 (2001)

[17] Soper S.A., Shera E.B., Davis L.M., Nutter H.L., Keller R.A., "The Photophysical Constants of Several Fluorescent Dyes Pertaining to Ultrasensitive Fluorescence Spectroscopy", Photochem. & Photobiol., **57**, 972-977 (1993)

[18] Magde D., Elson E., Webb W.W., "Thermodynamic Fluctuations in a Reacting System - Measurement By Fluorescence Correlation Spectroscopy", *Phys. Rev. Lett.*, **29**, 705-708 (1972)

[19] Jankowski T., Janka R., "ConfoCor 2 The Second Generation of Fluorescence Correlation Microscopes", in *Fluorescence Correlation Spectroscopy: Theory and Applications*, R. Rigler, E.S. Elson, Editors, Springer-Verlag Berlin Germany, pp. 331-345 (2001)

[20] Ball D.A., 'Single-Molecule Imaging With a Custom-Built Fluorescence Microscope', *MS Thesis*, University of Tennessee, Knoxville (2002)

[21] Davis L.M., Williams P.E., Cain H.M., Ball D.A., Parigger C.G., Matayoshi E.D., Swift K.M., "Comparison of Fluorescence Correlation Spectroscopy and Other Single-Molecule Data Analysis Methods for Assay of Protein-Ligand Interactions", *46-th Annual Meeting of the Biophysical Society*, San Francisco, California, February 23-27, 2002; *Biophysical J.*, **82**, 43a (2002).

[22] Kask P., Palo K., Fay N., Brand L., Mets Ü., Ullmann D., Jungmann J., Pschorr J., Gall K., "Two Dimensional Fluorescence Intensity Distribution Analysis: Theory and Applications", *Biophys. J.*, **78**, 1703-1713 (2000)

[23] Palo K., Mets Ü., Jäger S., Kask P., Gall K., "Fluorescence Intensity Multiple Distributions Analysis: Concurrent Determination of Diffusion Times and Molecular Brightness", *Biophys. J.*, **79**, 2858-2866 (2000)

[24] Wohland T., Rigler R., Vogel H., "The Standard Deviation In Fluorescence Correlation Spectroscopy", *Biophys. J.*, **80**, 2987-2999 (2001)

[25] Hess S.T., Webb W.W., 'Focal Volume Optics and Experimental Artifacts in Confocal Fluorescence Correlation Spectroscopy', *Biophys. J.*, **83**, 2300-2317 (2002)

[26] Magde D., Webb W.W., Elson E.L., "Fluorescence Correlation Spectroscopy. III. Uniform Translation and Laminar Flow", *Biopolymers*, **17**, 361-376 (1978)

[27] Kinjo M., Rigler R., "Ultrasensitive Hybridization Analysis Using Fluorescence Correlation Spectroscopy", *Nucleic Acids Research*, **23**, 1795-1799 (1995)

[28] Cooper G.M., in "The Cell: A Molecular Approach, Second Edition", ASM Press Washington D.C., U.S.A. (2000)

[29] Davis L.M., Ball D.A., *Private Communication*

[30] Gösch M., Blom H., Holm J., Heino T., Rigler R., "Hydrodynamic Flow Profiling in Microchannel Structures by Single Molecule Fluorescence Correlation Spectroscopy", *Anal. Chem.*, **72**, 3260-3265 (2000)

[31] Brinkmeier M., Rigler R., "Flow Analysis By Means of Fluorescence Correlation Spectroscopy", *Exp. Tech. Phys.*, **41**, 205-10 (1995)

[32] Brinkmeier M., Dörre K., Stephan J., Eigen M., "Two-Beam Cross-Correlation: A Method To Characterize Transport Phenomena in Micrometer-Sized Structures", *Anal. Chem.*, **71**, 609-616 (1999)

[33] Davis L.M., Williams P.E., Ball D.A., Swift K.M., Matayoshi E.D., "Data Reduction Methods for Application of Fluorescence Correlation Spectroscopy to Pharmaceutical Drug Discovery", *Curr. Pharm. Biotechnol.*, **4**, in press (2003)

[34] Constans A., "Beyond Sanger: Toward the $1,000 Genome", *The Scientist*, **Vol. 17 Issue 13**, 36-38, (2003)

[35] Braslavsky I., Hebert B., Kartalov E., Quake S.R., "Sequence Information Can Be Obtained From Single DNA Molecules", *Proc. Natl. Acad. Sci. USA*, **100**, 3960-3964 (2003)

[36] Koppel D.E., "Statistical Accuracy in Fluorescence Correlation Spectroscopy", *Phys. Rev. A*, **10**, 1938-1945 (1974)

[37] Qian H., "On the Statistics of Fluorescence Correlation Spectroscopy", *Biophys. Chem.*, **38**, 49-57 (1990)

[38] Saffarian S., Elson E.L., "Statistical Analysis of Fluorescence Correlation Spectroscopy: The Standard Deviation and Bias", *Biophys. J.*, **84**, 2030-2042 (2003)

[39] See webpage at http://www.picoquant.com/products/prt400.htm.

[40] See webpage at http://www.ni.com/labview.

Appendices

Appendix 1

MATLAB Code of Monte Carlo Simulation to Model Single-Molecule Imaging Experiments

```
%
%       diffuse1_5P - Models freely diffusing molecules
%
%       This program models the imaging of freely diffusing solution via a CCD camera.
%
%       The program assumes each molecule to be excited at each timestep, resulting
%       in a Poisson emission of photons.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% PARAMETERS                                                                    %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%
%       Simulation parameters
%

showcolorbar='yes';    % Whether or not to show the colorbar
pauseon='yes';         % Whether to pause after each final CCD frame

%
%       CCD parameters
%
nx_ccd=128;            % Number of CCD pixels in x-direction
ny_ccd=128;            % Number of CCD pixels in y-direction
dx_actual_ccd=13.0;  % Size of pixel in x-direction (microns)    22.5 for ICCD
dy_actual_ccd=13.0;  % Size of pixel in y-direction (microns)
frame_time=280e-3;   % Exposure time for frame (sec)           81.0 for ICCD
frame_transfer_time=1.52e-3; %Time for frame transfer (sec)
background_photons=0.01;    % (photons/pixel)
read_background=100.0;    % mean of the ccd readout noise (set to zer0 to subtract)
read_noise=36.0;       % variance of ccd readout noise
ADC_gain=1.8;          %

%       Flowcell parameters
%
flowcell_z=2.0;                % Depth of flowcell (microns)
```

```
%{1}
flow_speed=10000.0;      % Flow speed in microns/second

%
%       Optics parameters
%
%magnification=63;   % microscope magnification
magnification=157.5; %    with optivar
laser_power=325;                % Laser power (mW)
spot_x=125.0/1.2;       % laser spot length (microns)
spot_y=50.0/1.2;        % laser spot diameter (microns)
TIR_enhancement_factor=1; % intensity enhancement factor due to TIR (=3.7 at
d_ev=0.3, =1 at d_ev=inf)
NA=1.2;                         % Numerical aperture of objective lens
excitation_lambda=0.568;                % Wavelength of excitation light (microns)
% Evanescent field depth
d_ev=0.3;               % depth of evanescent field (microns)
spatial_resolution=0.08;     % timestep must be sufficiently small to track motion of
molecules (microns)
laser_exposure_time=3.0e-3;  % time that laser is on in seconds
Qdet=0.05;              % net detection collection efficiency
n=1.33;        % refractive index of water

% {2}
spot_per_defocus=2.0688333;
        % allow for image spot radius to change with defousing of object
        % calculated by Zeemax (NA dependent, magnification independent)
        % (microns radius per microns defocus)
zfocus = 1.0;
        % focal plane of objective lens relative to substrate surface (z=0)

%
%       Photophysics parameters
%       Molecule parameters
%
Dx=2.8e-6;              % Diffusion coefficient x-direction (cm^2/sec)
Dy=2.8e-6;              % Diffusion coefficient y-direction (cm^2/sec)
Dz=2.8e-6;              % Diffusion coefficient z-direction (cm^2/sec)
M=6.15e-10;                     % Molarity (moles/liter)
quench_factor=1;        % Quenching factor
%sigma_a=9.22e-15;  % absorption cross section (cm^2) This is much too large!
sigma_a=3.7e-16;        % absorption cross section (cm^2)
Qf=0.8;                 % fluorescence quantum efficiency
```

```
Qd=1.0e-5;              % photodestruction quantum efficiency
                        % (includes camera Qe)
emmission_lambda=0.61;          % Wavelength of emission light (microns)

% {3}
stick=0.0;
% Sticking parameter. Probability that a molecule will stick in hitting the z=0
%'wall' (0.0 = 0% chance of sticking, 1.0 = 100% chance etc.)
%
%       Detection parameters
%
first_thresh=20;        % First threshold for detection
second_thresh=60;       % Second threshold for detection

simulation_frames=6;
simulation_time=simulation_frames*(frame_time+frame_transfer_time);   % Total
simulation time (sec)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% INITIAL CALCULATIONS                                                  %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Set random number seed - can override for repeated experiment
%clocksum=sum(100*clock);
%rand('state',clocksum);
%disp(['Random state seed = ',num2str(clocksum)]);
rand('state',206647);           % Uncomment this line to use seed of your choice

dx_ccd=dx_actual_ccd/magnification;         % Size of pixel in x-direction (microns)
dy_ccd=dy_actual_ccd/magnification;         % Size of pixel in y-direction (microns)
if (spatial_resolution > dx_ccd) | (spatial_resolution > dy_ccd) | (spatial_resolution >
d_ev)
   warning('spatial_resolution is not small enough')
end

timestep=(spatial_resolution*0.0001)^2/2/Dx;        % Timestep for simulation (sec)
ensures sqrt(2D*timestep)=spatial_resolution
iterations_with_laser_on=laser_exposure_time/timestep;
iterations_with_laser_off= (frame_time+frame_transfer_time-
laser_exposure_time)/timestep;
timestep
iterations_with_laser_on
iterations_with_laser_off
```

```
%{1}
motion_per_iter=flow_speed * timestep;
        % to be worked out will be equivalent to ~ 1 micron per millisecond

% Diffusion movement
diff_fact_x=sqrt(2*Dx*timestep)*10000;
                        % Expected movement of molecule/timestep/dimension
diff_fact_y=sqrt(2*Dy*timestep)*10000;
                        % Expected movement of molecule/timestep/dimension
diff_fact_z=sqrt(2*Dz*timestep)*10000;
                        % Expected movement of molecule/timestep/dimension

% Calculate CCD size
ccd_x=nx_ccd*dx_ccd;                % size of ccd in x-direction (microns)
ccd_y=ny_ccd*dy_ccd;                % size of ccd in y-direction (microns)

% Calculate laser intensity in W/cm^2
laser_intensity=2.0*TIR_enhancement_factor*(laser_power/1000)/(pi*(spot_x/2/10000)*
(spot_y/2/10000));
laser_intensity

% Calculate energy of excitation photon and excitation rate
Egamma=(6.63e-34)*(3e8)/(excitation_lambda*1e-6);                % Joules
Ke=sigma_a*laser_intensity/Egamma/3;
                % excitations/second (factor of 3 for random orrientation wrt polarization)
Me=Ke*timestep;                % mean excitations/timestep (for a molecule at the surface)
Me

% Calculate normal standard deviation & variance which approximates Airy function
%       (see notes on 4/06/2000 in modelling notebook)
airy_sd=(0.4207*emmission_lambda/2/NA);
airy_var=airy_sd^2;

% {4}
% calculate point at which defocusing will become an issue
defocus_thresh=airy_sd/spot_per_defocus;

% Determine mean starting number of molecules to track
start_num=ccd_x*ccd_y*flowcell_z*M*(1e-15)*(6e23);
start_num
Nm=poiss_rnd(start_num,1);          % Generate a random number of molecules to track
Nm

% Starting x,y,z coordinates of molecules
```

```
m_x=rand(Nm,1)*ccd_x;
m_y=rand(Nm,1)*ccd_y;
m_z=rand(Nm,1)*flowcell_z;

cummulative_molecule_excitations=zeros(Nm,1);
excitations_until_bleached=-log(rand(Nm,1))/Qd;
QQ=Qf*Qdet/quench_factor;

% Initialize variables

%
% image matrix (expected number of frames ~ total time / ccd times)
%
if (exist('m_mat')), clear m_mat, end;
m_mat=zeros(ny_ccd,nx_ccd,ceil(simulation_time/(frame_time+frame_transfer_time)));


%figure;
%set(gcf,'Position',[10 200 1000 500]);
%set(gca,'Units','Pixels');
%set(gca,'Position',[15 15 980 435]);
%set(gca,'Xtick',[7]);
%set(gca,'Ytick',[7]);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%
% START LOOP TO MOVE MOLECULES                                  %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%
warning off                       % Stop the "empty=empty" warning

%time=timestep:timestep:simulation_time;   % time at each interval

for ccd_frames=1:simulation_frames
  m_mat(:,:,ccd_frames)=zeros(ny_ccd,nx_ccd);      % Clear the image

  %ccd_frames
  A=clock;
  A(1,4:5)        % used to measure time duration for each for each frame iteration

for i=1:iterations_with_laser_on

  %  {5}
  % Move molecules
  dx=diff_fact_x*randn(Nm,1);
```

dy=diff_fact_y*randn(Nm,1) **+ motion_per_iter; includes flow of water in y-direction;**
dz=diff_fact_z*randn(Nm,1);

% {6}
% Update molecule positions
m_x(find(m_z)) = m_x(find(m_z)) + dx(find(m_z));
m_y(find(m_z)) = m_y(find(m_z)) + dy(find(m_z));
m_z(find(m_z)) = m_z(find(m_z)) + dz(find(m_z));

% Fix entries outside flowcell
% {7}
% Let a certain amount of molecules stick to z=0
m_z(find((m_z<0)&(rand(length(m_z),1)<stick)))=0;
% Let the remainder of molecules leaving z-bounds bounce
m_z(find(m_z<0))=-m_z(find(m_z<0));
m_z(find(m_z>flowcell_z))=2*flowcell_z - m_z(find(m_z>flowcell_z));

% {8}
% Let the molecules leaving x and y-bounds wrap
m_x(find(m_x<0))=m_x(find(m_x<0))+ccd_x;
m_x(find(m_x>ccd_x))=m_x(find(m_x>ccd_x))-ccd_x;
m_y(find(m_y<0))=m_y(find(m_y<0))+ccd_y;
m_y(find(m_y>ccd_y))=m_y(find(m_y>ccd_y))-ccd_y;

% {9}
% Create Poisson excitations for molecules
Me_z=fill_place_vec(length(m_z),Me); case where penetration depth d_ev is infinite
%Me_z=exp(-m_z/d_ev)*Me;
molecule_excitations=poiss_mat_rnd(Me_z); % Array of number of excitations for each molecule for this iteration: mostly zero

% Find molecules that have been bleached and for those that are not, calculate number of photon
% excitations required for each molecule until bleaching occurs
bleached=find((molecule_excitations-excitations_until_bleached)>=0);
molecule_excitations(bleached)=excitations_until_bleached(bleached);
excitations_until_bleached=excitations_until_bleached-molecule_excitations;
excitations_until_bleached(bleached)=0;

% Calculate photons detected for each molecule
molecule_photons=poiss_mat_rnd(QQ*molecule_excitations); %This is incorrect; we could get more photons than excitations
                                                          %should be binomial random number

```
%   molecule_photons=binomial_rnd(QQ, molecule_excitations);

% Create photon vector based on molecule amplitudes and locations for focused (Airy
disc applicable)
% and defocused images of the molecules
%  {10}
airy_photon_x=fill_place_vec(molecule_photons.*(abs(m_z-
zfocus)<=defocus_thresh),m_x);
airy_photon_y=fill_place_vec(molecule_photons.*(abs(m_z-
zfocus)<=defocus_thresh),m_y);

%  {11}
defocus_photon_x=fill_place_vec(molecule_photons.*(abs(m_z-
zfocus)>defocus_thresh),m_x);
defocus_photon_y=fill_place_vec(molecule_photons.*(abs(m_z-
zfocus)>defocus_thresh),m_y);

% {12}
%defocused image variance for each molecule with m_z>defocus_thresh
defoc_var=fill_place_vec(molecule_photons.*(abs(m_z-
zfocus)>defocus_thresh),spot_per_defocus*abs(m_z-zfocus));

% 'Throw' photons into an Airy PSF based on Gaussian fit
%     (see notes on 4/6/2000 in modelling notebook)
a_x_airy=airy_photon_x+randn_mv(0,airy_var,length(airy_photon_x),1);
a_y_airy=airy_photon_y+randn_mv(0,airy_var,length(airy_photon_y),1);

% {13}
% 'Throw' photons into an out of focus image disk
[thrown_x,thrown_y] = rand_cir(defoc_var,length(defocus_photon_x),1);
a_x_defoc=defocus_photon_x+thrown_x;
a_y_defoc=defocus_photon_y+thrown_y;
% Concatentate the thrown photons into respective vectors
a_x=[a_x_airy; a_x_defoc];
a_y=[a_y_airy; a_y_defoc];

% Valid locations over CCD camera
valid_i=find((a_x>0) & (a_x<ccd_x) & (a_y>0) & (a_y<ccd_y));

% Assign valid locations to CCD pixels (bin)
pixel_x=ceil(a_x(valid_i)/dx_ccd);
pixel_y=ceil(a_y(valid_i)/dy_ccd);

% Bin into sparse matrix
s_mat=sparse(pixel_y,pixel_x,ones(length(valid_i),1),ny_ccd,nx_ccd);
```

```
  % Add to full matrix
  m_mat(:,:,ccd_frames) = m_mat(:,:,ccd_frames) + full(s_mat);
  %m_mat(:,:,ccd_frames)=full(s_mat);        % Uncomment if not adding to previous
frames

end;  %laser is now off


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
  % IMAGE ON SCREEN                                                          %
  %      do only at time of readout for 'fastest' simulation speeds       %

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 % Add in background photons and CCD ADC gain and readout noise
  noise_photons=poiss_rnd(background_photons*nx_ccd*ny_ccd,1);
  noise_smat=sparse(ceil(ny_ccd*rand(noise_photons,1)),...
    ceil(nx_ccd*rand(noise_photons,1)),ones(noise_photons,1),ny_ccd,nx_ccd);
  temp_mat = m_mat(:,:,ccd_frames)+full(noise_smat);       % add in background photons
  temp_mat = ADC_gain*temp_mat; % scale charges for ADC gain
  temp_mat = temp_mat + randn_mv(read_background,read_noise,ny_ccd,nx_ccd); %
add ADC noise, subtract offset
  temp_mat=(temp_mat-105.0)*63.0/45.0 +1.0;
  m_mat(:,:,ccd_frames)=temp_mat;

%   save 'i.mat' temp_mat
  image(m_mat(:,:,ccd_frames))   %display image of this frame
  jj=hsv(64);
  colormap([1,1,1;1,.8,.8;1,.7,.7; jj(1:59,:);zeros(2,3)]);
%      colormap(gray(Me*Qf*Qdet*2));
%   colormap(gray(Me*Qf*Qdet*3*ADC_gain));
  if (isequal(showcolorbar,'yes'))
    colorbar;
 end
  set(gca,'Xtick',[7]);
  set(gca,'Ytick',[7]);
  hold on
% show frame number
%   text(3,ny_ccd,sprintf('Step %d/%d     Time %d ms     Frame %d (readout)',...
%      i,length(time),round(time(i)*1000),ccd_frames),'Color',[0 0 0],...
%       'VerticalAlignment','baseline');
  drawnow;
```

```
    A=clock;
  A(1,4:5)              % calculates time at end of frame iteration

%   if (ccd_frames == 2)
%      stop              % used to stop simulation after one frame
     %   end

%   Now consider iterations when laser is off; molecules move, but no photophysics
 for i=1:iterations_with_laser_off

   % {5}
   % Move molecules
   dx=diff_fact_x*randn(Nm,1);
   dy=diff_fact_y*randn(Nm,1) + motion_per_iter;  includes flow of water in y-
direction;
   dz=diff_fact_z*randn(Nm,1);

   %  {6}
   % Update molecule positions
   m_x(find(m_z)) = m_x(find(m_z)) + dx(find(m_z));
   m_y(find(m_z)) = m_y(find(m_z)) + dy(find(m_z));
   m_z(find(m_z)) = m_z(find(m_z)) + dz(find(m_z));

   % Fix entries outside flowcell
   % {7}
   % Let a certain amount of molecules stick to z=0
   m_z(find((m_z<0)&(rand(length(m_z),1)<stick)))=0;
   % Let the remainder of molecules leaving z-bounds bounce
   m_z(find(m_z<0))=-m_z(find(m_z<0));
   m_z(find(m_z>flowcell_z))=2*flowcell_z - m_z(find(m_z>flowcell_z));

   % {8}
   % Let the molecules leaving x and y-bounds wrap
   m_x(find(m_x<0))=m_x(find(m_x<0))+ccd_x;
   m_x(find(m_x>ccd_x))=m_x(find(m_x>ccd_x))-ccd_x;
   m_y(find(m_y<0))=m_y(find(m_y<0))+ccd_y;
   m_y(find(m_y>ccd_y))=m_y(find(m_y>ccd_y))-ccd_y;

end     %End of iterations_with_laser_off

end      % End of simulation_frames

warning on                % Turn warning back on
```

Appendix 2

MATLAB Code to Run GUI to Perform Fits to ACF Data

```
function varargout = autocorrp1(varargin)
% AUTOCORRP1 Application M-file for autocorrp1.fig
%    FIG = AUTOCORRP1 launch autocorrp1 GUI.
%    AUTOCORRP1('callback_name', ...) invoke the named callback.

% Last Modified by GUIDE v2.0 07-Aug-2001 14:49:41

if nargin <= 1   % LAUNCH GUI
   if nargin == 0
      initial_dir = pwd;
   elseif nargin == 1 & exist(varargin{1},'dir')
      initial_dir = varargin{1};
   else
      errordlg('Input argument must be a valid directory',...
         'Input Argument Error!')
      return
   end

        fig = openfig(mfilename,'reuse');

        % Generate a structure of handles to pass to callbacks, and store it.
        handles = guihandles(fig);
        guidata(fig, handles);

   % Populate the listbox
   load_listbox(initial_dir,handles)

        if nargout > 0
                varargout{1} = fig;
        end

elseif ischar(varargin{1}) % INVOKE NAMED SUBFUNCTION OR CALLBACK

        try
                [varargout{1:nargout}] = feval(varargin{:}); % FEVAL switchyard
        catch
                disp(lasterr);
        end
```

```
end

% --------------------------------------------------------------------
% Read the current directory and sort the names
% --------------------------------------------------------------------
function load_listbox(dir_path,handles)

cd (dir_path)
path(path,dir_path)
dir_struct = dir(dir_path);
[sorted_names,sorted_index] = sortrows({dir_struct.name}');
handles.file_names = sorted_names;
handles.is_dir = [dir_struct.isdir];
handles.sorted_index = [sorted_index];
guidata(handles.figure1,handles)
set(handles.listbox1,'String',handles.file_names,...
        'Value',1)


%| ABOUT CALLBACKS:
%| GUIDE automatically appends subfunction prototypes to this file, and
%| sets objects' callback properties to call them through the FEVAL
%| switchyard above. This comment describes that mechanism.
%|
%| Each callback subfunction declaration has the following form:
%| <SUBFUNCTION_NAME>(H, EVENTDATA, HANDLES, VARARGIN)
%|
%| The subfunction name is composed using the object's Tag and the
%| callback type separated by '_', e.g. 'slider2_Callback',
%| 'figure1_CloseRequestFcn', 'axis1_ButtondownFcn'.
%|
%| H is the callback object's handle (obtained using GCBO).
%|
%| EVENTDATA is empty, but reserved for future use.
%|
%| HANDLES is a structure containing handles of components in GUI using
%| tags as fieldnames, e.g. handles.figure1, handles.slider2. This
%| structure is created at GUI startup using GUIHANDLES and stored in
%| the figure's application data using GUIDATA. A copy of the structure
%| is passed to each callback.  You can store additional information in
%| this structure at GUI startup, and you can change the structure
%| during callbacks.  Call guidata(h, handles) after changing your
%| copy to replace the stored original so that subsequent callbacks see
%| the updates. Type "help guihandles" and "help guidata" for more
```

```
%| information.
%|
%| VARARGIN contains any extra arguments you have passed to the
%| callback. Specify the extra arguments by editing the callback
%| property in the inspector. By default, GUIDE sets the property to:
%| <MFILENAME>('<SUBFUNCTION_NAME>', gcbo, [7], guidata(gcbo))
%| Add any extra arguments after the last argument, before the final
%| closing parenthesis.

% --------------------------------------------------------------
% Callback for list box - open .fig with guide, otherwise use open
% --------------------------------------------------------------

function varargout = listbox1_Callback(h, eventdata, handles, varargin)

global T
global GAC

if strcmp(get(handles.figure1,'SelectionType'),'open')
          index_selected = get(handles.listbox1,'Value');
          file_list = get(handles.listbox1,'String');
          filename = file_list{index_selected};
          if  handles.is_dir(handles.sorted_index(index_selected))
                    cd (filename)
                    load_listbox(pwd,handles)
          else
            [path,name,ext,ver] = fileparts(filename);
            switch ext
            case '.dat'
           create_data(filename,handles)
      case '.m'
          plot_data(filename,handles)
      otherwise
                    try
                    open(filename)
                    catch
                              errordlg(lasterr,'File Type Error','modal')
                    end
            end
  end
end

% --------------------------------------------------------------------

function varargout = file_menu_Callback(h, eventdata, handles, varargin)
```

```
% Stub for Callback of the uicontrol handles.file_menu.
if isempty(get(handles.axes1,'Children'))
   set(handles.print_plot_and_data,'Enable','off')
   set(handles.close_plot,'Enable','off')
else
   set(handles.print_plot_and_data,'Enable','on')
end

% --------------------------------------------------------------------

function varargout = create_new_plot_Callback(h, eventdata, handles, varargin)
% Stub for Callback of the uicontrol handles.create_new_plot.
      %run C program and print plot in window
[filename,pname] = uigetfile('*.dat','Select the binary .dat-file');
create_data(filename,pname,handles)
if filename==NULL
   quit cancel
end
filenamem=[filename '.m'];
plot_data(filenamem,handles)

% --------------------------------------------------------------------

function varargout = plot_data_Callback(h, eventdata, handles, varargin)
% Stub for Callback of the uicontrol handles.plot_data.
   % Get user input from GUI
[filename,pname] = uigetfile('*.m','Select the created M-file');
plot_data(filename,handles)

% --------------------------------------------------------------------

function varargout = print_plot_and_data_Callback(h, eventdata, handles, varargin)
% Stub for Callback of the uicontrol handles.print_plot_and_data.
print -f handles.figure1

% --------------------------------------------------------------------

function varargout = close_plot_Callback(h, eventdata, handles, varargin)
% Stub for Callback of the uicontrol handles.close_plot.
% --------------------------------------------------------------------

function varargout = exit_gui_Callback(h, eventdata, handles, varargin)
% Stub for Callback of the uicontrol handles.exit_gui.
delete(handles.figure1)
```

```
% --------------------------------------------------------------------
function varargout = fit_menu_Callback(h, eventdata, handles, varargin)
% Stub for Callback of the uicontrol handles.fit_menu.
if isempty(get(handles.axes1,'Children'))
   set(handles.gfit_one,'Enable','off')
else
   set(handles.gfit_one,'Enable','on')
end

set(handles.gfit_two,'Enable','off')

% --------------------------------------------------------------------
function varargout = gfit_one_Callback(h, eventdata, handles, varargin)
% Stub for Callback of the uicontrol handles.gfit_one.

global T
global GAC

set(handles.axes1,'nextplot','add')
half=(GAC(1)-1)/2 + 1;
it=1;
while GAC(it)>half, it=it+1; end
it
T(it)
a=[GAC(1)-1,(4^(1/3)-1)/T(it),.04];      %initial estimate for a
afirst=a
%semilogx(T,gfit(T,a(:)),'k','parent',handles.axes1)
[a(:),v]=fminsearch('ffit',a(:),[7],GAC,T)
semilogx(T,gfit(T,a(:)),'r','parent',handles.axes1)
%aa=[a,v];
set(handles.text6,'String','Fit Data');
var(1) = {'a1 = '};
var(2) = {'a2 = '};
var(3) = {'a3 = '};
var(5) = {'nu = '};
format short
str(1) = {a(1)};
str(2) = {a(2)};
str(3) = {a(3)};
str(5) = {v};
set(handles.text5,'String', var)
set(handles.text3,'String', str)
set(handles.axes1,'nextplot','replace');

% --------------------------------------------------------------------
```

```
function varargout = gfit_two_Callback(h, eventdata, handles, varargin)
% Stub for Callback of the uicontrol handles.gfit_two.

% --------------------------------------------------------------------
% --------------------------------------------------------------------

function plot_data(filename,handles)

        global T
        global GAC

        format short g

        filepath = pwd;
        addpath(filepath)

        fid=fopen(filename,'r');

%          DATA=fscanf(fid,'%g',[6 128]);    %from Simulation excel files
%         DATA=fscanf(fid,'%g',[3 142]);
        DATA=fscanf(fid,'%g',[4 4471]);   %for LI-COR Interim report files 5-15-01
%         DATA=fscanf(fid,'%g',[5 32768]);   %for LI-COR files
%         DATA=fscanf(fid,'%g',[5 32763]);   %for LI-COR files
                % after removing first 5 lines of data for which GAC ~=0
        status=fclose(fid);
        rmpath(filepath)

        axes(handles.axes1)

%         T=DATA(1,2:end); GAC=DATA(3,2:end);       %for Simulation excel files
%         T=DATA(1,:); GAC=DATA(4,:);        %for LI-COR files
        T=DATA(2,:); GAC=DATA(4,:);        %for LI-COR Interim report files 5-15-01
%         T=DATA(1,5:end); GAC=DATA(3,5:end);

        semilogx(T,GAC,'.','parent',handles.axes1)
        title=['Normalised Autocorrelation (' filename ')'];
        set(get(handles.axes1,'Xlabel'),'String','Time in seconds')
        set(get(handles.axes1,'Title'),'String',title)
        set(handles.axes1,'XGrid','on')
        set(handles.axes1,'YGrid','on')
        set(handles.text5,'String',' ')
        set(handles.text3,'String',' ')
        set(handles.text6,'String',' ')
        set(handles.axes1,'nextplot','replace');
% --------------------------------------------------------------------
```

```
function create_data(filename,pname,handles)

        global T
        global GAC
         [T,GAC]=acmatgui(filename,pname);
```

'gfit' function:

```
function v=gfit(t,a)

 [ldum,la]=size(t);

nnn=ones(ldum,la);

v=nnn +a(1)*(nnn +a(2)*t).^(-1) .* (nnn +a(2)*a(3)*t).^-0.5;
```

Appendix 3

Derivation of the ACF with Flow Included

Insertion of (12) and (34) into (9) gives

$$G_D^{(x)}(\tau) = \frac{1}{\sqrt{2\pi}\sqrt{2D\tau}}\left(\frac{2}{\pi\omega_0^2}\right)\int dx_1 \exp\left[-\frac{2x_1^2}{\omega_0^2}\right]\int dx_2 \exp\left[-\frac{2x_2^2}{\omega_0^2}\right]\exp\left[-\frac{(x_1 - x_2 - v\tau)^2}{4D\tau}\right],$$

for a beam radius $\omega_0 = 2\sigma$ and flow velocity v. Insertion of the expansion of $(x_1 - x_2 - v\tau)^2$ and equation (26) into the above gives

$$G_D^{(x)}(\tau) = \frac{1}{\sqrt{2\pi}\sqrt{2D\tau}}\left(\frac{2}{\pi\omega_0^2}\right)\exp\left[-\frac{v^2\tau}{4D}\right]\int dx_1 \exp\left\{-\left[\frac{1}{4D\tau}\left(1+\frac{2\tau}{\tau_D}\right)x_1^2 - \frac{v}{2D}x_1\right]\right\}$$

$$\int dx_2 \exp\left\{-\left[\frac{1}{4D\tau}\left(1+\frac{2\tau}{\tau_D}\right)x_2^2 - \left(\frac{x_1}{2D\tau} - \frac{v}{2D}\right)x_2\right]\right\}.$$

Integration over x_2 using equation (23) gives:

$$G_D^{(x)}(\tau) = \frac{1}{\sqrt{1+\frac{2\tau}{\tau_D}}}\left(\frac{2}{\pi\omega_0^2}\right)\exp\left[-\frac{v^2\tau}{4D}\right]$$

$$\int dx_1 \exp\left\{-\left[\frac{1}{4D\tau}\left(1+\frac{2\tau}{\tau_D}\right)x_1^2 - \frac{v}{2D}x_1\right]\right\}\exp\left\{\frac{1}{\left(1+\frac{2\tau}{\tau_D}\right)}\left[\frac{x_1^2}{4D\tau} - \frac{v}{2D}x_1 + \frac{v^2\tau}{4D}\right]\right\}$$

$$G_D^{(x)}(\tau) = \frac{1}{\sqrt{1+\frac{2\tau}{\tau_D}}}\left(\frac{2}{\pi\omega_0^2}\right)\exp\left[-\frac{v^2\tau^2}{2D\tau_D}\frac{1}{\left(1+\frac{2\tau}{\tau_D}\right)}\right]$$

$$\int dx_1 \exp\left\{-\left[\frac{1}{D\tau_D}\left(\frac{1+\frac{\tau}{\tau_D}}{1+\frac{2\tau}{\tau_D}}\right)x_1^2 - \frac{v\tau}{D\tau_D}\frac{1}{\left(1+\frac{2\tau}{\tau_D}\right)}x_1\right]\right\}.$$

Integration over x_1 yields:

$$G_D^{(x)}(\tau) = \left(\frac{2}{\pi\omega_0^2}\right)\frac{\sqrt{\pi D\tau_D}}{\sqrt{1+\frac{\tau}{\tau_D}}}\exp\left[-\frac{v^2\tau^2}{4D\tau_D}\frac{1}{\left(1+\frac{\tau}{\tau_D}\right)}\right]$$

$$= \frac{1}{\sqrt{\pi}\omega_0\sqrt{1+\frac{\tau}{\tau_D}}}\exp\left[-\frac{1}{\left(1+\frac{\tau}{\tau_D}\right)}\left(\frac{\tau}{\tau_F}\right)^2\right],$$

where the characteristic flow time τ_F is

$$\tau_F = \frac{\omega_0}{v} = \frac{\sqrt{4D\tau_D}}{v}.$$

Then in three dimensions, the normalized ACF with flow included is

$$G_D(\tau) = 1 + \frac{1}{N}\left[1+\frac{\tau}{\tau_D}\right]^{-1}\left[1+\left(\frac{\omega_0}{z_0}\right)^2\frac{\tau}{\tau_D}\right]^{-1/2}\exp\left[-\frac{1}{\left(1+\frac{\tau}{\tau_D}\right)}\left(\frac{\tau}{\tau_F}\right)^2\right].$$

Vita

Peter E. Williams was born in Swansea, Wales on July 26 1969. He was raised in nearby Llanelli, Wales until the age of 14 when his family moved to Bridgwater, Somerset, England. Driven by a life long passion for astronomy, he decided to return to education at the age of 25, beginning an MSci Astrophysics degree at Queen Mary College, University of London. His time at university was successful, gaining a First Class Honors degree and numerous awards for high academic achievements. He was also able to put what computer programming skills he had learnt in his spare time to good use building a computer model of circumstellar outflows from carbon stars. In the summer of 2001, Peter accepted a Graduate Research Assistantship at the University of Tennessee Space Institute under the supervision of Dr. Lloyd M. Davis. He is due to graduate from UTSI with a Master of Science degree in Physics in the fall of 2003.